

Tue Apr 6 07:57:22 2004

us-10-081-051-52.rng

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2004, 10:03:59 ; Search time 1689 Seconds
(without alignments)
11947.271 Million cell updates/sec

Title: US-10-081-051-52

Perfect score: 4750
Sequence: 1 gatccctgctaactctttag9.....ccacatctgctaactgac 4750

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

N Geneseq_29Jan04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002s:*
7: Geneseqn2003as:*
8: Geneseqn2003bs:*
9: Geneseqn2003cs:*
10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4750	100.0	4750	6	AAD48255 Ehrlichia
2	1416	29.8	1416	6	AAD48257 Ehrlichia
3	597	12.6	597	6	AAD48258 Ehrlichia
4	371	7.8	371	6	AAD48256 Ehrlichia
5	260.4	5.5	1370	6	ABK97886 DNA encod
6	187.6	5.9	1017	6	ACA28500 Prokaryot
7	186.4	3.9	8056	7	ABZ10246 Haematopo
8	175.2	3.7	1029	7	ACA29483 Prokaryot
9	170.6	3.6	996	7	ACA47630 Prokaryot
10	168.6	3.5	996	7	ACF75135 Staphyloc
11	168.2	3.5	999	7	AAS55337 Staphyloc
12	168.2	3.5	999	7	ACA27525 Prokaryot
13	165.8	3.5	110000	2	AAK20248_00
14	165.8	3.5	110000	2	AAK20248_01
15	165.4	3.5	999	7	ACA46288 Prokaryot
16	162.2	3.4	8056	7	ABZ10246 Haematopo
17	160	3.4	1557	8	ADA31919 DNA encod
18	159.6	3.4	83391	6	ABQ67093 Human ang
19	155.2	3.3	7814	4	AAS46530 Tumour su
20	155	3.3	1065	7	ACA33900 Prokaryot
21	155	3.3	1071	6	ADG90723 E. faeciu
22	155	3.3	5641	9	ABL33396 Human imm
23	154.2	3.2	918	2	AAV74824 Staphyloc

C	24	153	3.2	15548	6	ABL34155 Human imm
C	25	150.6	3.2	8056	7	ABZ10100 Haematopo
C	26	148.8	3.1	6314	6	ABL54314 Chemical
C	27	148.4	3.1	1020	7	ACA18468 Prokaryot
C	28	148.4	3.1	1023	4	AAS5118 Enterococ
C	29	147.2	3.1	1026	4	AAS51287 Enterococ
C	30	147.2	3.1	6132	6	ABL32863 Human imm
C	31	146.2	3.1	110000	6	ABQ69245_20
C	32	146.2	3.1	319633	6	ABQ67194 Listeria
C	33	145.4	3.0	11745	6	ABK28332 DNA trans
C	34	144.2	3.0	7597	6	ABL33013 Human imm
C	35	143	3.0	47108	6	ABK31511 Human ang
C	36	141	3.0	5611	6	ABQ67070 Human ang
C	37	139.8	2.9	1062	9	ADC91772 E. faeciu
C	38	139.8	2.9	2791	6	ABQ70805 Listeria
C	39	139	2.9	6175	6	ABL33307 Human imm
C	40	138.8	2.9	37515	6	ABQ66998 Human ang
C	41	137	2.9	18154	6	ABL32254 Human imm
C	42	136.8	2.9	1008	4	AAS53331 Haemophil
C	43	136.8	2.9	1008	7	ACA34182 Prokaryot
C	44	136.8	2.9	110000	2	AAT42063_06
C	45	136.2	2.9	11691	6	ABL34240 Human imm

ALIGNMENTS

RESULT 1	AD48255	ABZ10100
ID	AD48255 standard; DNA; 4750 BP.	
XX		
AC	AD48255:	
XX		
DT	24-FEB-2003 (first entry)	
XX		
DE	Ehrlichia ruminantium 19hwr02 protein encoding DNA.	
XX		
XX	Vaccine; immunity; rickettsial infection; spotted fever; heart water;	
KW	typhus; pathogen; immunostimulant; glycerol 3-phosphate dehydrogenase;	
KW	19hwr02; antibacterial; enzyme; gene; ds.	
XX		
OS	Ehrlichia ruminantium.	
XX		
PH	key	Location/Qualifiers
FT	CDS	632..1615
FT		/*tag= a
FT		/product= "Hypothetical glycerol 3-phosphate
FT		dehydrogenase (19hwr02)"
FT		/note= "This region is specifically claimed as SEQ ID NO:
FT		54 in claim 1 of the specification"
XX		
XX	MO20026652-A2.	
XX		
XX	29-AUG-2002.	
PD		
XX		
XX	20-FEB-2002; 2002WC-US005772.	
PF		
XX		
XX	20-FEB-2001; 2001US-0269944P.	
PR		
XX		
XX	(UYFL) UNIV FLORIDA.	
PA		
XX	Barbet AF, Whitmore WW, Kamper SM, Simbi BH, Ganta RR,	
XX	Moreland AL, Wangi DM, Mogueire TC, Mahan SM;	
PI		
XX	WPI: 2002-723186/78.	
DR		
XX	P-PSDB; AAE31069.	
XX		
XX	New Ehrlichia ruminantium polynucleotides, useful as vaccines for	
PT	inducing protective immunity, and protecting animals or humans against	
PT	rickettsial diseases, e.g. typhus, spotted fever or heart water.	
XX		
PS	Claim 1, Page 131-133; 206pp; English.	

Ds	1081	TATAGACTGCTGGTGAATAATAAGAACTTGGGAAATCATGATGAAACAAATAGTA	1144
Qy	1141	TGATGTTCTAAAAATATATACCATCAAGATATTATAGGTGACAGATTGAGCTGCATT	1200
Ds	1141	TGATGTTCTAAAAATATATACCATCAAGATATTATAGGTGACAGATTGAGCTGCATT	1200
Qy	1201	AAGAACAATATGGCAATGGCATGTGGAAATATCGCTGGAAAAAATTGGATATATGC	1266
Ds	1201	AAGAACAATATGGCAATGGCATGTGGAAATATCGCTGGAAAAAATTGGATATATGC	1266
Qy	1261	TGTTGCTACTGTTATTAAGCTAAGGCGATGAATGAATTAACATATATATAGCAAAAA	1320
Ds	1261	TGTTGCTACTGTTATTAAGCTAAGGCGATGAATGAATTAACATATATATAGCAAAAA	1320
Qy	1321	TCATTCATTAAGTCTTCAATATTAATTTGGTGCATCATGCTTGGGATCTTAATTTAAC	1386
Ds	1321	TCATTCATTAAGTCTTCAATATTAATTTGGTGCATCATGCTTGGGATCTTAATTTAAC	1386
Qy	1381	ATGTACAAACAGACATTCACGCCAATATGCTTTTGGCATAGAAATAGAAAAAGTAGAAA	1444
Ds	1381	ATGTACAAACAGACATTCACGCCAATATGCTTTTGGCATAGAAATAGAAAAAGTAGAAA	1444
Qy	1441	TATTAATCATTTATAGTATCAACAACTTAAGCTTTTGAAGGAACCATGATCTGTAACC	1500
Ds	1441	TATTAATCATTTATAGTATCAACAACTTAAGCTTTTGAAGGAACCATGATCTGTAACC	1500
Qy	1501	ACTGATATCATTTAGCAAAAACTTAATGTAGAACTACCAATTTGCATATCTATTACA	1566
Ds	1501	ACTGATATCATTTAGCAAAAACTTAATGTAGAACTACCAATTTGCATATCTATTACA	1566
Qy	1561	TTTATTTAATGAGATATATCATAGTAAAGCCATATCAAAATTTATCTTAGTCTAT	1622
Ds	1561	TTTATTTAATGAGATATATCATAGTAAAGCCATATCAAAATTTATCTTAGTCTAT	1622
Qy	1621	CATCCTTGATTAATTTCAACAGATGATAATTAAGCGCAAAAAATATGATATACAGT	1688
Ds	1621	CATCCTTGATTAATTTCAACAGATGATAATTAAGCGCAAAAAATATGATATACAGT	1688
Qy	1681	TCACACATATATAAAAAGAGTCAATACCCATCGTTAACTCCAACTTAAGTATGTAACT	1740
Ds	1681	TCACACATATATAAAAAGAGTCAATACCCATCGTTAACTCCAACTTAAGTATGTAACT	1740
Qy	1741	AATATACATAACAACCTTAAGTATACCTTAATTTCCAGTCAATTAATATATTAACAAT	1800
Ds	1741	AATATACATAACAACCTTAAGTATACCTTAATTTCCAGTCAATTAATATATTAACAAT	1800
Qy	1801	CGTATCTCTATCTCAATTAACCAATTTGATCAGATTAATAATAGTCAATTTAACTT	1866
Ds	1801	CGTATCTCTCTCAATTAACCAATTTGATCAGATTAATAATAGTCAATTTAACTT	1866
Qy	1861	TATCTATTTTAAATAATATATGTTGGCACTTTTCCATCAATCAATGATATATATGA	1922
Ds	1861	TATCTATTTTAAATAATATATGTTGGCACTTTTCCATCAATCAATGATATATATGA	1922

QY	1921	CATATATTTACATATAAAATTAACATAATTAAGAAAGCAACATACATATATTTAAATTTACCTAAAC	1980
Dp	1921	CATATATTTACATATAAAATTAACATAATTAAGAAAGCAACATACATATATTTAAATTTACCTAAAC	1980
QY	1981	ACCTATTTACACGTAATATTTACATAGATGTCCTGACATATAATTAATTTATCATCATCTTTGA	2040
Dp	1981	ACCTATTTACACGTAATATTTACATAGATGTCCTGACATATAATTAATTTATCATCATCTTTGA	2040
QY	2041	CATATTTCAAAACATATAAATATGATPAACTCGTTTTTTTTTGCCCATTAATCTTATATATAA	2100
Dp	2041	CATATTTCAAAACATATAAATATGATPAACTCGTTTTTTTTTGCCCATTAATCTTATATATAA	2100
QY	2101	TTGCTAAATTACCCAAAAGGTTAAATAACATTTCTGCAACAAATGAGAGATACAA	2160
Dp	2101	TTGCTAAATTACCCAAAAGGTTAAATAACATTTCTGCAACAAATGAGAGATACAA	2160
QY	2161	TACACGTCAAAATTAATTAACAAATTTCTAAATATACACGCTACCTTAACATPAGAGAAC	2220
Dp	2161	TACACGTCAAAATTAATTAACAAATTTCTAAATATACACGCTACCTTAACATPAGAGAAC	2220
QY	2221	AATAGTATATCATCTCTAATGATAGTATCCCTACAGCATGTATCATAAATTAAGAAAA	2280
Dp	2221	AATAGTATATCATCTCTAATGATAGTATCCCTACAGCATGTATCATAAATTAAGAAAA	2280
QY	2281	TGCTATATCATATTTATTCAGATTTCCAAATTTCCAACTGCAACAAACAACTCAAAGTTC	2340
Dp	2281	TGCTATATCATATTTATTCAGATTTCCAAATTTCCAACTGCAACAAACAACTCAAAGTTC	2340
QY	2341	TATATATAACAATTCAGAAATTAATATATATCTTCCAAACACATAGTATCAAAATTTCTATA	2400
Dp	2341	TATATATAACAATTCAGAAATTAATATATATCTTCCAAACACATAGTATCAAAATTTCTATA	2400
QY	2401	ATACATCATATCTTTTAAAAAGTGACTCACAAATTCATACATPAAATATTTATPACATG	2460
Dp	2401	ATACATCATATCTTTTAAAAAGTGACTCACAAATTCATACATPAAATATTTATPACATG	2460
QY	2461	CAATGTTTATTCAGAAACATTACACAACATCCAGAAATTTTCAATTTACGGTAAATGCTGA	2520
Dp	2461	CAATGTTTATTCAGAAACATTACACAACATCCAGAAATTTTCAATTTACGGTAAATGCTGA	2520
QY	2521	ACGTTTCCATTCGCAACAAATTCAGAGATGCGGATACATTTACTCTCCAAACGAGATPACTG	2580
Dp	2521	ACGTTTCCATTCGCAACAAATTCAGAGATGCGGATACATTTACTCTCCAAACGAGATPACTG	2580
QY	2581	AGTTTTTAAATACTGTCCTTTCAAAATCTCTCAAGACTTTACGTAATGGTACAGAAATAC	2640
Dp	2581	AGTTTTTAAATACTGTCCTTTCAAAATCTCTCAAGACTTTACGTAATGGTACAGAAATAC	2640
QY	2641	TTTGAACCTTAAACATCATATTAATAGGAGATTATGATCTATATATCTATATGTAATCTTT	2700
Dp	2641	TTTGAACCTTAAACATCATATTAATAGGAGATTATGATCTATATATCTATATGTAATCTTT	2700
QY	2701	TGCTGTAATCATCTTCCTTTAGAGAAATTCATATAATTAATTAATTCATTTATATACTTACGTA	2760
Dp	2701	TGCTGTAATCATCTTCCTTTAGAGAAATTCATATAATTAATTAATTCATTTATATACTTACGTA	2760
QY	2761	TTTGCTTAATAGTAACTCGGCACTCATATAGATCTGATTTGATTAAGCTTATCATCACTTAA	2820
Dp	2761	TTTGCTTAATAGTAACTCGGCACTCATATAGATCTGATTTGATTAAGCTTATCATCACTTAA	2820
QY	2821	TACATGAGTACACAAACCTATTTTTTTCACATGCTATTTCAATAATTCCTTACACAATTC	2880
Dp	2821	TACATGAGTACACAAACCTATTTTTTTCACATGCTATTTCAATAATTCCTTACACAATTC	2880
QY	2881	CGGATATCTGACACAAATTTCTACTTAAAGAGAGGATCTTAATTTGGAAGACATTTAACT	2940
Dp	2881	CGGATATCTGACACAAATTTCTACTTAAAGAGAGGATCTTAATTTGGAAGACATTTAACT	2940
QY	2941	ATAATATATAATCTCACAAAACCTACAGCTTTTACTTCACCTTCAATATCTTTGGAAGA	3000
Dp	2941	ATAATATATAATCTCACAAAACCTACAGCTTTTACTTCACCTTCAATATCTTTGGAAGA	3000

QY	3001	AGACACAAATAATCTCAATCTCAATACGAAAGGAACTGTAATTTCCCTATATATTT	3065
Db	3001	AGACACAAATAATCTCAATCTCAATACGAAATCTTAATTTCCCTATATATTT	3065
QY	3061	TCCCTCTGTATATATCTGAGTAATCTTAATATGCGATGCAATCTGTACTTC	3120
Db	3061	TCCCTCTGTATATATCTGAGTAATCTTAATATGCGATGCAATCTGTACTTC	3120
QY	3121	ATCATATAATAGGACCAATGATTTGGTTGCTCATATTTCCAAATAGAGGACTCT	3180
Db	3121	ATCATATAATAGGACCAATGATTTGGTTGCTCATATTTCCAAATAGAGGACTCT	3180
QY	3181	ATGAGACAATATATATTAATCTTTCCTCACTACCAATATATTAACAAATATATTAAGC	3240
Db	3181	ATGAGACAATATATATTAATCTTTCCTCACTACCAAAATATATTAACAAATATATTAAGC	3240
QY	3241	TGGTAGCATGATGATAGACATAGATATTAATATGAGATGATACATCCCTTGATTTTATG	3300
Db	3241	TGGTAGCATGATGATAGACATAGATATTAATATGAGATGATACATCCCTTGATTTTATG	3300
QY	3301	TATTAAGCTTAGCAACTACTTCTTTCCACAACCTGCGAAACAGTAATAGATAGACT	3360
Db	3301	TATTAAGCTTAGCAACTACTTCTTTCCACAACCTGGGAAACAGTAATAGATAGACT	3360
QY	3361	CGATGTAGTAGCGCTTATTAATACACTTCTCAATTAAGATACAGGGAGTTAC	3420
Db	3361	CGATGTAGTAGCGCTTATTAATACACTTCTCAATTAAGATACAGGGAGTTAC	3420
QY	3421	GACTATTTCAATATCTCTCAATATGCTGATTTCAACTCATCATTTTCTTACGTAATCTAAC	3480
Db	3421	GACTATTTCAATATCTCTCAATATGCTGATTTCAACTCATCATTTTCTTACGTAATCTAAC	3480
QY	3481	AGACTCATATAGCTCTCTTTCAACTACTTAAATCTCTGCTGTAAGAGGCTTTCTAT	3540
Db	3481	AGACTCATATAGCTCTCTTTCAACTACTTAAATCTCTGCTGTAAGAGGCTTTCTAT	3540
QY	3541	ATAATCATATAAGCAACCATATATGACAGAGCTTTACAGCATGCGCAATATTAACATGCGCACT	3600
Db	3541	ATAATCATATAAGCAACCATATATGACAGAGCTTTACAGCATGCGCAATATTAACATGCGCACT	3600
QY	3601	AATCATATATAAGGGCAATAATAGATATCTTTCTTAAGCTTTCCAGTACACTTAATCC	3660
Db	3601	AATCATATATAAGGGCAATAATAGATATCTTTCTTAAGCTTTCCAGTACACTTAATCC	3660
QY	3661	ATCAATATCAAGATCTCTTAAACAATATATTCOAATATATACAAATCAGGCTCTTTTCATA	3720
Db	3661	ATCAATATCAAGATCTCTTAAACAATATATTCOAATATATATACAAATCAGGCTCTTTTCATA	3720
QY	3721	AGGCATCTGTATGCGCGGATATAACATCATATCGCTAAATTTAGTGCATAATTAATCATCACT	3780
Db	3721	AGGCATCTGTATGCGCGGATATAACATCATATCGCTAAATTTAGTGCATAATTAATCATCACT	3780
QY	3781	TAAATATATCTTTTATATATGATTTCTGATATCAACTCATCATCAACAATAATCTTACAG	3840
Db	3781	TAAATATATCTTTTATATATGATTTCTGATATCAACTCATCATCAACAATAATCTTACAG	3840
QY	3841	AATATACAATCTTTCCTTGACANTTCAAAATCCTGCGCATATATTAACATATATCTAA	3900
Db	3841	AATATACAATCTTTCCTTGACANTTCAAAATCCTGCGCATATATTAACATATATCTAA	3900
QY	3901	TGAATATCTGCAACACAAATGTAATCTTTAAAGTACACTCATATATATACACACACACA	3960
Db	3901	TGAATATCTGCAACACAAATGTAATCTTTAAAGTACACTCATATATATACACACACACA	3960
QY	3961	CTAATGATCAACAACCAATATATTAATTAATTAATTTTACTTATTAATTTTAAACATAT	4020
Db	3961	CTAATGATCAACAACCAATATATTAATTAATTAATTTTACTTATTAATTTTAAACATAT	4020
QY	4021	TATATAAATATATGTTATCTATATTAATCTAATTTATACAGCATATAGTTTCTATATATCCAAA	4080
Db	4021	TATATAAATATATGTTATCTATATTAATCTAATTTATACAGCATATAGTTTCTATATATCCAAA	4080
QY	4081	AAGAAAAATATAAACACTTATATAAAAAAATATATCATCATCTTGTAAACCTATATAGTACT	4140

Db	4081	AAAAAAAAATAAAAACCTTATATAAAAAATTATATATCAACTTTGTAAACCTATATGTAACT	4140
Qy	4141	AATTTATTAATTAATTTATTTCTATTAATTAACCTTCCCTGGAACACGCATGCTATAGCTTT	4200
Db	4141	AATATTTAATTAATTAATTTATTTCTATTAATTAACCTTCCCTGGAACACGCATGCTATAGCTTT	4200
Qy	4201	TCCATACATAAAAACTCACTTGAACCTTTATGATATTCATTAATTAATTCACGCCAACACAGAT	4260
Db	4201	TCCATACATAAAAACTCACTTGAACCTTTATGATATTCATTAATTAATTCACGCCAACACAGAT	4260
Qy	4261	AAGATCATTAATTATGGTAAATTTTATTAATTAACCACTAGACAAATTAATCCCACTAAT	4320
Db	4261	AAGATCATTAATTATGGTAAATTTTATTAATTAACCACTAGACAAATTAATCCCACTAAT	4320
Qy	4321	GACATCAACATATGATATGATATGAACCAATACCAACCAACAGAGATTTATCATTTGACTA	4380
Db	4321	GACATCAACATATGATATGATATGAACCAATACCAACCAACAGAGATTTATCATTTGACTA	4380
Qy	4381	CTTACGCTATTAATCTATGATGAGTTAAACGACCATCTCCATGTTATGAAGTCAATCAT	4440
Db	4381	CTTACGCTATTAATCTATGATGAGTTAAACGACCATCTCCATGTTATGAAGTCAATCAT	4440
Qy	4441	CCCGTATATATACATTAATATCAACATTAATATGTCACAAACTATCAATGATAGAAATTTTGT	4500
Db	4441	CCCGTATATATACATTAATATCAACATTAATATGTCACAAACTATCAATGATAGAAATTTTGT	4500
Qy	4501	TATATATATACCAATTTTGCAAAAGCAGATATTTCTTGACCTGTTATCTGTAAGTATTTG	4560
Db	4501	TATATATATACCAATTTTGCAAAAGCAGATATTTCTTGACCTGTTATCTGTAAGTATTTG	4560
Qy	4561	GTAAACACCTCTTACTAGGCACTTTTGTATCCATGATGACATGACTATATTTGTTTC	4620
Db	4561	GTAAACACCTCTTACTAGGCACTTTTGTATCCATGATGACATGACTATATTTGTTTC	4620
Qy	4621	TCAATATGAGCCAAAGGCAAAAAAGATGACCTTGCAATCAAGATTTTACGTAATTTTCAT	4680
Db	4621	TCAATATGAGCCAAAGGCAAAAAAGATGACCTTGCAATCAAGATTTTACGTAATTTTCAT	4680
Qy	4681	CAGAACTTACATATTCATTAACCATCAATCAATATTTTATCAACAGTAAATCCCAATTTGA	4740
Db	4681	CAGAACTTACATATTCATTAACCATCAATCAATATTTTATCAACAGTAAATCCCAATTTGA	4740
Qy	4741	CTAATTTGATC 4750	
Db	4741	CTAATTTGATC 4750	
RESULT 2			
AAd48257			
XX	ID	AAd48257 standard; DNA; 1416 BP.	
XX	AC	AAd48257;	
XX	DT	24-FEB-2003 (first entry)	
XX	DE	Ehrlichia ruminantium 19hwerf3 protein encoding DNA.	
XX	XX	Vaccine; immunity; rickettsial infection; spotted fever; heart water;	
XX	XX	typhus; pathogen; immunostimulant; antibacterial; 19hwerf3; enzyme; gene;	
XX	XX	nitrogen assimilation regulatory protein; ds.	
XX	OS	Ehrlichia ruminantium.	
XX	XX	MO20026652-A2.	
XX	PD	29-AUG-2002.	
XX	PE	20-FEB-2002; 2002MO-US005772.	
XX	PR	20-FEB-2001; 2001US-0269944P.	
XX	PA	(UFL) UNIV FLORIDA.	

XX	Babbel AP, Whitmore MW, Kamper SM, Simdi BH, Ganta RR;
PI	Moroland AL, Mwangi DM, McGuire IC, Mahan SM;
DR	WPI; 2002-723186/79.
XX	New Ehrlichia ruminantium poly nucleotides, useful as vaccines for
PT	inducing protective immunity, and protecting animals or humans against
PT	rickettsial diseases, e.g. typhus, spotted fever or heart water.
XX	
PS	Claim 1; Page 137-138; 206pp; English.
CC	The present invention relates to nucleic acid vaccines for conferring
CC	immunity to rickettsial infection, including Ehrlichia ruminantium
CC	(formerly Cowdria ruminantium). The invention also relates to novel E.
CC	ruminantium polynucleotides and their corresponding proteins. Sequences
CC	of the invention are useful for inducing immunity, particularly
CC	protective immunity. They are also useful for detecting the presence of
CC	E. ruminantium in a biological sample. They are useful in vaccines for
CC	protecting animals or humans against rickettsial diseases, e.g. typhus,
CC	spotted fever or heart water. Sequences of the invention are useful for
CC	detecting antibodies to pathogens. The present sequence is a DNA encoding
CC	E. ruminantium nitrogen assimilation regulatory protein (19mworf3)
XX	
SQ	Sequence 1416 BP; 472 A; 297 C; 161 G; 486 T; 0 U; 0 Other;
	Query Match 29.8%; Score 1416; DB 6; Length 1416;
	Best Local Similarity 100.0%; Pred. No.7.3e-215; Indels 0; Gaps 0;
	Matches 1416; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	2467 TTATTTCAGAAACATTAACAATCCAGAAATTTTCAATTTACGGTGTAATGCTGAAGCTTC 2526
DB	1 TTATTTCAGAAACATTAACAATCCAGAAATTTTCAATTTACGGTGTAATGCTGAAGCTTC 60
QY	2527 CATTCCAAACAAATTCAGACAGTGTGTATATTAATACCCTCGAAAACGAGTAACTAGAGTTTT 2586
DB	61 CATTCCAAACAAATTCAGACAGTGTGTATATTAATACCCTCGAAAACGAGTAACTAGAGTTTT 120
QY	2587 TAAATACTGTCCTTCCAATTTCTTTCAGAGCTTTAAGTAATGTTACAGAAATAACTTTAGC 2646
DB	121 TAAATACTGTCCTTCCAATTTCTTTCAGAGCTTTAAGTAATGTTACAGAAATAACTTTAGC 180
QY	2647 ACTTAAACATCATTAATAGGCGAATTAAGTATCATATCTACAGGTAAATCTTTGCTGT 2706
DB	181 ACTTAAACATCATTAATAGGCGAATTAAGTATCATATCTACAGGTAAATCTTTGCTGT 240
QY	2707 AATCATCTCCTTAGAGATTTCAATTAATTAATTCATTATTAACATTAAGCTAATTTGCT 2766
DB	241 AATCATCTCCTTAGAGATTTCAATTAATTAATTCATTATTAACATTAAGCTAATTTGCT 300
QY	2767 TAATCTACCGGCATTCATATGACGTCGATTTGGATTAAGGCTTCATCACCTTAATACATG 2826
DB	301 TAATCTACCGGCATTCATATGACGTCGATTTGGATTAAGGCTTCATCACCTTAATACATG 360
QY	2827 AGTACAGAAACCTATTTTTTTTTCAGATGCTATTCAATAAATCCTACACAAATTCGGTAT 2886
DB	361 AGTACAGAAACCTATTTTTTTTTCAGATGCTATTCAATAAATCCTACACAAATTCGGTAT 420
QY	2887 ATCTGTCAATATCTTACTAAGAAGCGTACTGCTAATGGAAGGACATTTATCATATATA 2946
DB	421 ATCTGTCAATATCTTACTAAGAAGCGTACTGCTAATGGAAGGACATTTATCATATATA 480
QY	2947 TAAATCTCAGAAACCTACCAAGCTTTTACTTCATTCAATATCTTTGGAGAAGACAC 3006
DB	481 TAAATCTCAGAAACCTACCAAGCTTTTACTTCATTCAATATCTTTGGAGAAGACAC 540
QY	3007 AATATATCTGACATCTATACCAAGAGATCTTACATTTTCCCTATATATTTTCCCTC 3066
DB	541 AATATATCTGACATCTATACCAAGAGATCTTACATTTTCCCTATATATTTTCCCTC 600
QY	3067 CTGTATATATCTAGTATCTTAATTCGGTATCTGATATCGTAATCTGTACTATCTAT 3126
DB	601 CTGTATATATCTAGTATCTTAATTCGGTATCTGATATCGTAATCTGTACTATCTAT 660

QY 3127 AATAACGACATGATTTGCTGCTATTAATCCAAATATGAGAGAGTACTGATGAGA 3186
 DB 661 AATAACGACATGATTTGCTGCTATTAATCCAAATATGAGAGAGTACTGATGAGA 720
 QY 3187 CAATATATATTAATCTTCTCCTACATCAAAATATATTAACCAATATATTAATGCTGATG 3246
 DB 721 CAATATATATTAATCTTCTCCTACATCAAAATATATTAACCAATATATTAATGCTGATG 780
 QY 3247 CATGATGAGTACATATATATTAATGAGATGATCAATCCCTGATTTTATATATGATG 3306
 DB 781 CATGATGAGTACATATATATTAATGAGATGATCAATCCCTGATTTTATATGATGATG 840
 QY 3307 CCTGCAACTACTCTCTTCCAAACAGCTGGGCAACAGATATAGATATGATGATGATGATG 3366
 DB 841 CCTGCAACTACTCTCTTCCAAACAGCTGGGCAACAGATATAGATATGATGATGATGATG 900
 QY 3367 AGTAGCTGCTTTATTAATCACTTCTCAAAATTAAGATATGAGAGAGTATACCGATAT 3426
 DB 901 AGTAGCTGCTTTATTAATCACTTCTCAAAATTAAGATATGAGAGAGTATACCGATAT 960
 QY 3427 TTCAATATCTCAATATGCTGATTTTCAATGATTTTCTCTCAATATCTACCAAGCTG 3486
 DB 961 TTCAATATCTCAATATGCTGATTTTCAATGATTTTCTCTCAATATCTACCAAGCTG 1020
 QY 3487 TATAGCTCTCTTACCAATCACTTATCTCTCTGTAAGAGCTTTTCTATATATG 3546
 DB 1021 TATAGCTCTCTTACCAATCACTTATCTCTCTGTAAGAGCTTTTCTATATATG 1080
 QY 3547 ATAAGCAACCCATATGAGAGAGCTTTACAGAGATGAGCAATTTACATGCCACTATCAT 3606
 DB 1081 ATAAGCAACCCATATGAGAGAGCTTTACAGAGATGAGCAATTTACATGCCACTATCAT 1140
 QY 3607 AATAACGAGCAATATGAGATGATCTTTTCAAGCTTTTCAAGCACTTATCATCAT 3666
 DB 1141 AATAACGAGCAATATGAGATGATCTTTTCAAGCTTTTCAAGCACTTATCATCAT 1200
 QY 3667 ATCAAGATCTCTTACCAATATGATCAATATATGATCAATCAATGCTTTTCAATAGCAT 3726
 DB 1201 ATCAAGATCTCTTACCAATATGATCAATATATGATCAATCAATGCTTTTCAATAGCAT 1260
 QY 3727 CTGATGCGGATATTAACATCAATGCTATTTATGATGATATATATCATCATCAT 3786
 DB 1261 CTGATGCGGATATTAACATCAATGCTATTTATGATGATATATATCATCATCAT 1320
 QY 3787 ATCTTTTATATGATTTCTGATATCAATCTTATCATCAATCAATCAATCAATATAT 3846
 DB 1321 ATCTTTTATATGATTTCTGATATCAATCTTATCATCAATCAATCAATCAATATAT 1380
 QY 3847 CAATCTTTCTTGAAGATTTCAAAATCCCTGCTCAT 3882
 DB 1381 CAATCTTTCTTGAAGATTTCAAAATCCCTGCTCAT 1416
 RESULT 3
 AAD48258
 ID AAD48258 standard; DNA; 597 BP.
 AC AAD48258;
 DT 24-FEB-2003 (first entry)
 DE Ehrlichia ruminantium 19hwoFF41 protein encoding DNA.
 EE
 XX
 KM Vaccine; immunity; rickettsial infection; spotted fever; heart water;
 KM typhus; pathogen; immunostimulant; antibacterial; 19hwoFF41; gene;
 KM cell division protein; ftsQ; ds.
 OS Ehrlichia ruminantium.
 XX
 XX
 PN NC020026652-A2.
 XX
 PD 29-AUG-2002.

XX 20-FEB-2002; 2002WO-US005772.
 PF 20-FEB-2001; 2001US-0269944P.
 PR (UFL) UNIV FLORIDA.
 PA Barbet AF, Whitmore WM, Kamper SM, Simbi BH, Ganta RR;
 PI Moreland AL, Whang DM, McGuire TC, Mahan SM;
 DR WPI; 2002-723186/78.
 XX
 PT New Ehrlichia ruminantium polynucleotides, useful as vaccines for
 PT inducing protective immunity, and protecting animals or humans against
 PT rickettsial diseases, e.g. typhus, spotted fever or heart water.
 XX
 PS Claim 1, Page 138-139; 206pp; English.
 CC The present invention relates to nucleic acid vaccines for conferring
 CC immunity to rickettsial infection, including Ehrlichia ruminantium
 CC (formerly Cowdria ruminantium). The invention also relates to novel E.
 CC ruminantium polynucleotides and their corresponding proteins. Sequences
 CC of the invention are useful for inducing immunity, particularly
 CC protective immunity. They are also useful for detecting the presence of
 CC E. ruminantium in a biological sample. They are useful in vaccines for
 CC protecting animals or humans against rickettsial diseases, e.g. typhus,
 CC spotted fever or heart water. Sequences of the invention are useful for
 CC detecting antibodies to pathogens. The present sequence is a DNA encoding
 CC E. ruminantium hypothetical cell division protein (ftsQ, 19hwoFF41)
 XX
 SQ Sequence 597 BP; 220 A; 117 C; 68 G; 192 T; 0 U; 0 Other;
 Query Match 12.6%; Score 597; DB 6; Length 597;
 Best Local Similarity 100.0%; Pred. No. 1,9e-85;
 Matches 597; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4154 TTATTTCTATATTAATTAACCTTCTGGAACAGCATGCTATAGTTTCCATATCAAAA 4213
 DB 1 TTATTTCTATATTAATTAACCTTCTGGAACAGCATGCTATAGTTTCCATATCAAAA 60
 QY 4214 CTCACCTTAACCTTATATGATATATCAATATATGAGCAAGCAGCATATGATATAT 4273
 DB 61 CTCACCTTAACCTTATATGATATATCAATATATGAGCAAGCAGCATATGATATAT 120
 QY 4274 AGGTATATTTATATTAACCATAGCAATATATATATATATATATATATATATAT 4333
 DB 121 AGGTATATTTATATTAACCATAGCAATATATATATATATATATATATATATAT 180
 QY 4334 TGTGATGATGAAACCATACCAAGCAAGATATATATATATATATATATATATATAT 4393
 DB 181 TGTGATGATGAAACCATACCAAGCAAGATATATATATATATATATATATATATAT 240
 QY 4394 ATCTAGATGATTAACGACCATCTCCATGATATGAGATCAATCATCCGATATATTA 4453
 DB 241 ATCTAGATGATTAACGACCATCTCCATGATATGAGATCAATCATCCGATATATTA 300
 QY 4454 ATATATCAAT 4513
 DB 301 ATATATCAAT 360
 QY 4514 ATTGCAAAAGCAGATATCTTGAACCTGATATCTGTAAGATATTTGTAACAGCTCTT 4573
 DB 361 ATTGCAAAAGCAGATATCTTGAACCTGATATCTGTAAGATATTTGTAACAGCTCTT 420
 QY 4574 AACTGAGGCAATTTTGAATCCATGATGATGATGATGATGATGATGATGATGATG 4633
 DB 421 AACTGAGGCAATTTTGAATCCATGATGATGATGATGATGATGATGATGATGATG 480
 QY 4634 AGGCACAAAAAGATAGACCTTGATCAACAGTTTATGATTTTATATATATATATAT 4693
 DB 481 AGGCACAAAAAGATAGACCTTGATCAACAGTTTATGATTTTATATATATATATAT 540
 QY 4694 TTCAATACCATCAATCAATATTTTATCAACAGTAAATCCCAATGATGATATGATC 4750

Db 541 TTGATTCACATCAATTTTATTCACAGTAATCCAAATTCAGTAAATTCATC 597

RESULT 4

AAD48256 standard; DNA; 371 BP.

XX AAD48256;

XX 24-FEB-2003 (first entry)

XX Ehrlichia ruminantium 19hwoxfli protein encoding DNA.

XX Vaccine; immunity; rickettsial infection; spotted fever; heart water;

XX typhus; pathogen; immunostimulant; integrase; recombinase; 19hwoxfli;

XX antibacterial; enzyme; gene; ds.

XX Ehrlichia ruminantium.

XX WO200266652-A2.

XX 29-AUG-2002.

XX 20-FEB-2002; 2002WO-US005772.

XX 20-FEB-2001; 2001US-0269944P.

XX (UYFL) UNIV FLORIDA.

XX Barbet AF, Whitmore MW, Kamper SM, Simbi BH, Ganta RR,

XX Moreland AL, Mwambi DM, McGuire TC, Mahan SM,

XX WPI; 2002-723186/78.

XX New Ehrlichia ruminantium polynucleotides, useful as vaccines for

XX inducing protective immunity, and protecting animals or humans against

XX rickettsial diseases, e.g. typhus, spotted fever or heart water.

XX Claim 1; Page 134; 206pp; English.

XX The present invention relates to nucleic acid vaccines for conferring

XX immunity to rickettsial infection, including Ehrlichia ruminantium

XX (formerly Cowdria ruminantium). The invention also relates to novel E.

XX ruminantium polynucleotides and their corresponding proteins. Sequences

XX of the invention are useful for inducing immunity, particularly

XX E. ruminantium in a biological sample. They are useful in vaccines for

XX protecting animals or humans against rickettsial diseases, e.g. typhus,

XX spotted fever or heart water. Sequences of the invention are useful for

XX detecting antibodies to pathogens. The present sequence is a DNA encoding

XX protein

XX

XX

XX

Db 181 CTCGAATCTCCAAATTTTAAACATGACAAAGATTATTATATGTAACAGAACCCCTGACTC 240

Qy 241 TTATGCAAGAACTCTATPAACTTATCAAGGCTCTTACATPAAAGAACTGATTTCAA 300

Db 241 TTATGCAAGAACTCTATPAACTTATCAAGGCTCTTACATPAAAGAACTGATTTCAA 300

Qy 301 GAATACCTCTTCTCTTGTTCATCCATATTTTCCAAATTTATPACATATATGAGTCT 360

Db 301 GAATACCTCTTCTCTTGTTCATCCATATTTTCCAAATTTATPACATATATGAGTCT 360

Qy 361 TTATTCGACAT 371

Db 361 TTATTCGACAT 371

XX ABR97886/c

XX ABR97886 standard; DNA; 1370 BP.

XX ABR97886;

XX 10-OCT-2002 (first entry)

XX DNA encoding Lawsonia intracellularis antigenic ntrc protein.

XX Immunogen; fibs; filr; ntrc; gluh; motb; clyc; ytfm; ytfn; porcine;

XX pig; avian; bird; porcine proliferative enteropathy; PPE;

XX intestinal adenomatosis complex; porcine intestinal adenomatosis; PIA;

XX necrotic enteritis; proliferative haemorrhagic enteropathy;

XX regional ileitis; haemorrhagic bowel syndrome; vaccine; antibacterial;

XX porcine proliferative enteritis; Campylobacter spp.-induced enteritis;

XX gene; ds.

XX Lawsonia intracellularis.

XX WO200238594-A1.

XX 16-MAY-2002.

XX 09-NOV-2001; 2001WO-AU001462.

XX 10-NOV-2000; 2000AU-00001381.

XX 17-NOV-2000; 2000US-0249596P.

XX (AGRI-) AGRIC VICTORIA SERVICES PTY LTD.

XX (AUPO-) AUSTRALIAN PORK LTD.

XX (PFI2) PFIZER PROD INC.

XX Rosey EL, King KW, Good RT, Strugnell RA;

XX WPI; 2002-557448/59.

XX P-PSDB; ABG68912.

XX New immunogenic polypeptide comprising epitope of Lawsonia spp.

XX polypeptide such as fibh, filr, ntrc, gluh, motb, polypeptides, useful in

XX vaccines for treatment of porcine proliferative enteropathy in pigs and

XX birds.

Db 181 CTCGAATCTCCAAATTTTAAACATGACAAAGATTATTATATGTAACAGAACCCCTGACTC 240

Qy 241 TTATGCAAGAACTCTATPAACTTATCAAGGCTCTTACATPAAAGAACTGATTTCAA 300

Db 241 TTATGCAAGAACTCTATPAACTTATCAAGGCTCTTACATPAAAGAACTGATTTCAA 300

Qy 301 GAATACCTCTTCTCTTGTTCATCCATATTTTCCAAATTTATPACATATATGAGTCT 360

Db 301 GAATACCTCTTCTCTTGTTCATCCATATTTTCCAAATTTATPACATATATGAGTCT 360

Qy 361 TTATTCGACAT 371

Db 361 TTATTCGACAT 371

XX ABR97886/c

XX ABR97886 standard; DNA; 1370 BP.

XX ABR97886;

XX 10-OCT-2002 (first entry)

XX DNA encoding Lawsonia intracellularis antigenic ntrc protein.

XX Immunogen; fibh; filr; ntrc; gluh; motb; clyc; ytfm; ytfn; porcine;

XX pig; avian; bird; porcine proliferative enteropathy; PPE;

XX intestinal adenomatosis complex; porcine intestinal adenomatosis; PIA;

XX necrotic enteritis; proliferative haemorrhagic enteropathy;

XX regional ileitis; haemorrhagic bowel syndrome; vaccine; antibacterial;

XX porcine proliferative enteritis; Campylobacter spp.-induced enteritis;

XX gene; ds.

XX Lawsonia intracellularis.

XX WO200238594-A1.

XX 16-MAY-2002.

XX 09-NOV-2001; 2001WO-AU001462.

XX 10-NOV-2000; 2000AU-00001381.

XX 17-NOV-2000; 2000US-0249596P.

XX (AGRI-) AGRIC VICTORIA SERVICES PTY LTD.

XX (AUPO-) AUSTRALIAN PORK LTD.

XX (PFI2) PFIZER PROD INC.

XX Rosey EL, King KW, Good RT, Strugnell RA;

XX WPI; 2002-557448/59.

XX P-PSDB; ABG68912.

XX New immunogenic polypeptide comprising epitope of Lawsonia spp.

XX polypeptide such as fibh, filr, ntrc, gluh, motb, polypeptides, useful in

XX vaccines for treatment of porcine proliferative enteropathy in pigs and

XX birds.

Db 181 CTCGAATCTCCAAATTTTAAACATGACAAAGATTATTATATGTAACAGAACCCCTGACTC 240

Qy 241 TTATGCAAGAACTCTATPAACTTATCAAGGCTCTTACATPAAAGAACTGATTTCAA 300

Db 241 TTATGCAAGAACTCTATPAACTTATCAAGGCTCTTACATPAAAGAACTGATTTCAA 300

Qy 301 GAATACCTCTTCTCTTGTTCATCCATATTTTCCAAATTTATPACATATATGAGTCT 360

Db 301 GAATACCTCTTCTCTTGTTCATCCATATTTTCCAAATTTATPACATATATGAGTCT 360

Qy 361 TTATTCGACAT 371

Db 361 TTATTCGACAT 371

XX ABR97886/c

XX ABR97886 standard; DNA; 1370 BP.

XX ABR97886;

XX 10-OCT-2002 (first entry)

XX DNA encoding Lawsonia intracellularis antigenic ntrc protein.

XX Immunogen; fibh; filr; ntrc; gluh; motb; clyc; ytfm; ytfn; porcine;

XX pig; avian; bird; porcine proliferative enteropathy; PPE;

XX intestinal adenomatosis complex; porcine intestinal adenomatosis; PIA;

XX necrotic enteritis; proliferative haemorrhagic enteropathy;

XX regional ileitis; haemorrhagic bowel syndrome; vaccine; antibacterial;

XX porcine proliferative enteritis; Campylobacter spp.-induced enteritis;

XX gene; ds.

XX Lawsonia intracellularis.

XX WO200238594-A1.

XX 16-MAY-2002.

XX 09-NOV-2001; 2001WO-AU001462.

XX 10-NOV-2000; 2000AU-00001381.

XX 17-NOV-2000; 2000US-0249596P.

XX (AGRI-) AGRIC VICTORIA SERVICES PTY LTD.

XX (AUPO-) AUSTRALIAN PORK LTD.

XX (PFI2) PFIZER PROD INC.

XX Rosey EL, King KW, Good RT, Strugnell RA;

XX WPI; 2002-557448/59.

XX P-PSDB; ABG68912.

XX New immunogenic polypeptide comprising epitope of Lawsonia spp.

XX polypeptide such as fibh, filr, ntrc, gluh, motb, polypeptides, useful in

XX vaccines for treatment of porcine proliferative enteropathy in pigs and

XX birds.

CC avian animal subject. (1) is preferably useful for vaccinating porcine
 CC animals against intestinal diseases collectively known as porcine
 CC proliferative enteropathy (PPE), previously known as intestinal
 CC adenomatous complex, porcine intestinal adenomatosis (PIA), necrotic
 CC enteritis, proliferative haemorrhagic enteropathy, regional ileitis,
 CC haemorrhagic bowel syndrome, porcine proliferative enteritis and
 CC Campylobacter spp.-induced enteritis. (1) is also useful in vaccines for
 CC the prophylaxis and treatment of PPE in birds. This sequence encodes a
 CC Lawsonia intracellularis immunogenic peptide used in the creation of a
 CC porcine vaccine described in the invention
 CC
 XX

Sequence 1370 BP; 443 A; 201 C; 278 G; 448 T; 0 U; 0 Other;

Query Match 5.5%; Score 260.4; DB 6; Length 1370;
 Best Local Similarity 52.1%; Pred. No. 3.2e-32;
 Matches 708; Conservative 0; Mismatches 631; Indels 19; Gaps 5;

QY 2482 ACAGATCCAGATTTTCATTTAGCGTGAATGCTGAACGTTCCATTCCACCAATTC 2541
 DB 1358 ATTAATACATAGCTTTTACCTTTCTATATTAATTAATCTTTCAAGTCAATAGCTTC 1299
 QY 2542 AGCAGTTCGTATCATTTACCTCCAAACGAGATTAAGTATTTTAAATAGTCTTC 2601
 DB 1298 TGTATACGGGTATATTTCTTGTATAGCATGTAATTTTCAGTTAAATTTTGTTTC 1239
 QY 2602 AAATTTTCAGACTTACGTAATGCTACA---GAATTAATTTAGCACTTAAACATC 2658
 DB 1238 AAAAGCTATTTTACCGGTAAATCTATATCGGGAAGAAAGTTATAGGGAAGAA 1179
 QY 2659 ATTATAGCGCAATTAATATCTATATCTAGCTGAATCTTTGCTGATATCATCTCCTT 2718
 DB 1178 TTCTTGGTAAATTTTAAATTTGCTTTTACCTTATAGATCTGTCAACATCTTCTT 1119
 QY 2719 AGGAGATTTCAATTAATTAATTCATTTAGCAATTTAGTATTTAGTTACCTG 2778
 DB 1118 CCCGTAATTAAGATTAACCAATGCTTCTACAAATTAATTAATTTCTTATCTTCCCTG 1059
 QY 2779 CCATTATAGCACTGCAATGCTATTTAAGCTCATACATTAATCAATGAGTACAAACC 2838
 DB 1058 CCACATATTTGTTCAATACAGGAGACCTCTCTAATAAATTAAGCGCTCAGCT 999
 QY 2839 TATTTTTCAGATGCTATTCATTAATTAATTAATTAATTAATTAATTAATTAATTA 2898
 DB 998 ATAACTTACTCAATGATCAACCAATTAATTAATTAATTAATTAATTAATTAATTA 939
 QY 2899 TTCTACAAAGAGGTACTCTAATTTGAGAGCAATTTAATCTATTAATTAATTTCTTCA 2958
 DB 938 TTACGAGAGGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 879
 QY 2959 AAACCTACGAGCTTTTACTTCAATCTTTGGAAGAGACAAATTAATTTCTAC 3018
 DB 878 AAATGTTCCATCGTATAGCGCTTCAAGATTTCTTATTTGCTGCAATTAATCTTCA 819
 QY 3019 ATCTATCTACAGGAATCTTACTATTTTCCCTATATATTTTCCCTCTGTAATATCT 3078
 DB 818 ATCAACTTATATGTTTCACTACCAATTTTTCAAACATTTGTTTGGCAATTAAG 759
 QY 3079 GAGTAATCTTAATGCTATGCTATGTAATCTGTAATCTGTAATCTTAATTAATTAAGTAC 3138
 DB 758 CAATTTTTCGTTGCTTTTAAATCTCAATCTCTATTTCTCAAGAAATTAATTTTCC 699
 QY 3139 ATGATTTGCTGCTCTATTAATTCATATGAGAGGTAATCTTAATGACATTAATTAAT 3198
 DB 698 TTATATGCGCACTCAAAAGACCTGACAGAGAGATGCGCACGTAAGGCC----- 644
 QY 3199 ACTTCTGCTACGCAAAATATTAACAGATTAATTAATTAATTAATTAATTAATTAATTA 3258
 DB 643 -CTTTTTCAGCAAAATGTTGCTTTCAATCAATTTCTTCAAGGATTAAGCAAAATTT 585
 QY 3259 CATGATTAATTAATGAGTATCAATCCCTTGATTTTATGATTAATGCTTAAGCAATTA 3318
 DB 584 AAGAGCTATAATGTTTGTATATGATGAGCTTCTTGTGCAATGCTTGTGCACTTA 525

QY 3319 TTCCTTCCAAACCTGGGCAACCGTATGATATAGTATGATGATGATGATGATGATGAT 3378
 DB 524 CTCTTACCTGTATCAATTTCTCTCTGTAAGTATGATGATGATGATGATGATGATGAT 465
 QY 3379 ATTAATCAATCTTCAATTAATGATTAACAGAGGAGTTACCGAATTAATTAATTTCTC 3438
 DB 464 TGATATTAATCTTAAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 405
 QY 3439 A---AATGCTATTTCAATCATATTTTCTTCTATGATATTTACAGACTTATAGCTCT 3495
 DB 404 AGTATATCAAGTATGATTAATTTGTTTCCCTTTTATCTTATCTGTTTCTATAGCTCT 345
 QY 3496 CTTTCAACTAATCTTATCTTCTCTGTAAGGCTTTTCTTATTAATTAATTAATTAATTA 3555
 DB 344 ATTACCTGTATTAATTAAGACTTTTTCAGAAAGAGGCTTTTCAATTAATTAATTAAT 285
 QY 3556 CATATGCAAGACTTTTAACAGAGTGGCAATTAATTAATTAATTAATTAATTAATTA 3615
 DB 284 TTGAGGATATGCTGTATACAGAGCTTCAATTTGCTGATGATGATGATGATGATGAT 225
 QY 3616 CAATTAAGGATTAATCTTTTATAGCTTTTCCAGTACACTTATGATGATGATGATGAT 3675
 DB 224 TTAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 169
 QY 3676 TCTTAACCATATATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3735
 DB 168 -GAAGCCAAATATCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 111
 QY 3736 GGATTAACCATCACTGCTATTTATGATGATTAATTAATTAATTAATTAATTAATTA 3795
 DB 110 TTCTTCTGCTGAAGCTCTTTTAAATCTTCAATGAGCCCTTCAATTAATTAATTTCTT 51
 QY 3796 TAGATTTCTGATATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3833
 DB 50 T-GAATCTAATTAAGTCTTCAATCAATTAATTAATTAATTAATTAATTAATTAATTA 14

RESULT 6

ACA28500
 ID ACA28500 standard; DNA; 1017 BP.

AC ACA28500;

DT 19-JUN-2003 (first entry)

DE Prokaryotic essential gene #10157.

KW Antisense; ds; prokaryotic essential gene; cell proliferation;

KW drug design; gene.

OS Clostridium botulinum.

PN WO200277183-A2.

PD 03-OCT-2002.

PF 21-MAR-2002; 2002WO-US009107.

PR 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

PA (ELIT-) ELITRA PHARM INC.

PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 DR WPI; 2003-029926/02.
 DR P-PSDB; ABU24630.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening

PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
PS
XX Claim 14; SEQ ID NO 16370; 1766bp; English.

CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pcc_sequences

XX
XX Sequence 1017 BP; 412 A; 106 C; 229 G; 270 T; 0 U; 0 Other;
SQ
Query Match 3.9%; Score 187.6; DB 7; Length 1017;
Best Local Similarity 52.1%; Pred. No. 1e-20;
Matches 499; Conservative 0; Mismatches 444; Indels 15; Gaps 3;

QY 635 AAAATCAGTATTGTTGAGTCAGAGATCTTGGCAGCAGCAATGCACTGTCAGCA 694
DB 31 AAGAGTTGTTTGTGAGCAGAGTATTGTTGAGCAGCTTTAGCTGATCTGAGAAA 90
QY 695 CATGGTATCAGTAACTTATGGGAGCGTATGCTAGAGAAATATACATATAAAGCT 754
DB 91 AAGAGTTGAAAGTTACATATGCGATAGAAATATACATATAGTAGAGATTAATATC 150
QY 755 TACCGAAAAATTTAAATATTTACCAATATCATCTTACAGACACATATATGCAAC 814
DB 151 AAAAAAGAAATATATAATATTACCTAAGTAGTATACCTGAGGGGTAAAGCTTAT 210
QY 815 AGCAATATAGCAAGATATTATCTGACACATACATGATATTATCTTACATCTTCA 874
DB 211 AATGAATATGAAAAAGTAT---GACATATGTAATTTATATTTAGTACGATCTTCC 267
QY 875 CAACAATTACGACCATATGTACACAATATCAACAACAAGCATATGTATTAATACT 934
DB 268 CATGTGATATAGCAGGTATGCAAAAAAGATAC-----AACCTTTATATAAGAAATCAA 321
QY 935 CCAATATTAATTTAGTAAAGTATGCAAAATTAATCATCACTCAATTTCCAGAGAAATA 994
DB 322 ATATATGTATAGTATAGCTTAAAGGTATAGAGAGGTGACGAGAAAGTTATCAGAGTCC 381
QY 995 GCAGAGAAATTTTCAATATATATCAATTTTATATCTCTGTGCAAGTTTGTCTAAA 1054
DB 382 ATTAAGAAAGAACTGCTTAATATCTCTATATAGTATATGTCAGGCCCTTACGACAGAG 441
QY 1055 GAATTTGAGAACATCTTCTGTTAGTATAGTACTGCTGGTGAATAAAGAACTTGGT 1114

DB 442 GAAGTAGCTCAAGATATACCTTACACAGTAGTGGCTTTCAGAGCATATGATATAGCT 501
QY 1115 GAATCATTTAGTAAAGAAATTAAGTATGATGTTCTTAAATATATATACCATCAAGATAT 1174
DB 502 GAAAGATTCAGAGACCTTTTATGACTATTAATTTAGAGTGTATATGATGATTTA 561
QY 1175 ATGCGTATCAGATTTGAGCTGATTAAGAAATATATGCAATTCATGATGGAATATC 1234
DB 562 GTGAGGTGAAATATAGAGCGCTGTAAATATATATGCTTATGCTGAGAGGCTATCT 621
QY 1235 GCTGAAAAAATTTAGTATATATGCTGTGCTACTGTTATTAATCAATTAAGCATGATGA 1294
DB 622 GATGATATAGTATATGAGATATACCAAGGACGCTTATATGCAAGAGTATGATGA 681
QY 1295 ATTAAGACATATATATGCAAAAAATCATTTCAATATGATCTTCAATTAATGCTCA 1354
DB 682 ATATATAGA-----ATAGTCAAAAGCTTGAAGGTAAACAGAAACCTTTTCAGGTCT 735
QY 1355 TCATGCTTGGAGATCTAATATTAATTAATGATGACACAGAACTTCAGCAATATGCTTT 1414
DB 736 ACAGTATGAGATTTAATATGATATGATATGATATGATATGATATGATATGATATG 795
QY 1415 GGACTAGAAATAGGAAAGGTAGAAATATTAATTAATATGATCAACAACCTTAAGCTT 1474
DB 796 GGTATCTATATCGTTAAAGGCTATATGACAGAGAAAGCAATAGAAAGGTATGATGA 855
QY 1475 GTTGAAGGAAACCGATCTGTTAAACCGATATATCTTATGCAAAAAAATTAATATGAGA 1534
DB 856 GTGAGGAAATTAAGAACTTGTAAAGCTTTTATATCTTAAAGAAAGCTATATGAGA 915
QY 1535 CTACCAATTTGATATCTATTTTAACTTTATACATGAGCAATATATCACTAATTAAG 1592
DB 916 ATCCCTATTAACAGATCATTAATATTAAGTTATTTGAAAAATTAAGAGCTATATATG 973

RESULT 7
ABZ10246/c
ID ABZ10246 standard; DNA; 8056 BP.
XX
AC ABZ10246;
XX
DT 16-JUN-2003 (first entry)
XX
DE Haematopoietic cell proliferation disorder related DNA sequence #386.
XX
XX Human; haematopoietic cell proliferation disorder; cytostatic;
XX gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;
XX cytosine methylation state; gene; ds.
XX
OS Homo sapiens.
XX
PN W0200277272-A2.
XX
PD 03-OCT-2002.
XX
PE 26-MAR-2002; 2002WO-BE003401.
XX
PR 26-MAR-2001; 2001US-0278333P.
XX
PA (EPIC-) EPIDENOMICS AG.
XX
PI Berlin K, Braun A, Distler J, Guetig D, Howe A, Mueller J;
PI Olek A, Piepenbrock C, Adorjan P, Grabs G, Lesche R, Liu E;
PI Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T, Pellet C;
PI Schwöpe I, Ziebarth H;
XX
DR WPI; 2003-018942/01.
XX
PT Detecting and differentiating between haematopoietic cell proliferative
PT disorders, comprises contacting a target nucleic acid with a reagent that
PT distinguishes between methylated and non-methylated CpG dinucleotides.
XX
XX Claim 28; SEQ ID NO 386; 117bp; English.

XX The present invention describes a method for detecting and
CC differentiating between haematopoietic cell proliferative disorders
CC associated with at least 1 gene and/or their regulatory regions in a
CC subject. The method comprises contacting a target nucleic acid in a
CC biological sample obtained from the subject with at least 1 reagent,
CC which distinguishes between methylated and non-methylated CpG
CC dinucleotides within the target nucleic acid. AB209861 to AB211118
CC represent specifically claimed nucleotide sequences from the present
CC invention. Oligonucleotides from the present invention can be used: for
CC differentiating between healthy haematopoietic cells and proliferative
CC disorder haematopoietic cells; for differentiating between acute
CC lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
CC determining the cytosine methylation state and/or single nucleotide
CC polymorphisms (SNPs) of haematopoietic cell proliferation disorder
CC related sequences and their complements; and as primers for the
CC amplification of haematopoietic cell proliferation disorder related DNA
CC sequences. The nucleotide sequences from the present invention can also
CC be used for detecting a predisposition to, differentiation between
CC subclasses, diagnosis, prognosis, treatment and/or monitoring of
CC haematopoietic cell proliferative disorders. The present method enables a
CC highly specific classification of haematopoietic cell proliferative
CC disorders allowing for improved and informed treatment of patients
CC
XX

Sequence 8056 BP, 3711 A; 0 C; 371 G; 3974 T; 0 U; 0 Other;

Query Match 3.9%; Score 186.4; DB 7; Length 8056;
Best Local Similarity 43.8%; Pred. No. 1.7e-20;
Matches 1900; Conservative 0; Mismatches 2341; Indels 99; Gaps 21;

QY 308 TCTCTCTGTCATCCATATTTTCCATTTATTTTCAATATCATAGAGCTTTTATTCG 367
DB 4643 TTTAAATTTTTTAAATAATTAATTAATTTTAAATAATTAATTAATTTTAA
QY 368 ACATACATATTAATAATCAATCAAAACAATATAGTACTATCTATTATTAATAATC 427
DB 4585 ATTTAAATTTTAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4526
QY 428 CTCATATGCTTTAGCAAGACTCAATATTAAGACACTTAACTTATTAATTAATTT 487
DB 4525 AAACAAATAATTAATTAATTTTATTAATTAATTAATTAATTAATTAATTAATTT 4466
QY 488 CAATATTTATTAATCTCATCATATTTTATTAATTAATTAATTAATTAATTAATTT 547
DB 4465 CATATTTATTTTATTTTATTTTATTCATCAT--TTTCAATTTATTTTCACTTAAT 4408
QY 548 AGTACACATTTCTTAACACTAGATATATTAATTAATTAATTAATTTTGACTGT 607
DB 4407 ACATAAATAATTAATTAATTAATTTTAAATTTTAAATTTTATTTTATTA 4348
QY 608 TTTCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTTGAAGC 667
DB 4347 AAATTAATTTCAAAAAAATTAATTAATTAATTTTAAATTAATTTTCAATTTAA 4288
QY 668 ACAGCAATAGCAATGCTGCTGACAGCAATGATATCAATTAATTTGAGGAGCTAT 727
DB 4287 ATACAAAAAATAATTTTATTAATTAATTTTATTAATTTTATTAATTAATTAAT 4228
QY 728 CA----TAAATATTTACATTAATTAATTAATTAATTAATTTTATTAATTTTCCAC 783
DB 4227 TAAATTTCAATTAATTTTATTTTATTAATTAATTAATTAATTAATTAATTAAT 4168
QY 784 ATATCATCTACAGCAATATATGCAACGACATATAGCAAGATTAATTTGCAAA 843
DB 4167 AAAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTT 4108
QY 844 CAATATATGATTAATTTTAACTATTTCTTACAGCAATTTAGCAGCAATATGCAAT 903
DB 4107 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTT 4048
QY 904 ACAACACAAGCAATATGTAATAAATCTCAATATTAATTTTATTAAGTAATGCA 963
DB 4047 ATACAAAAA-----AATTTGACACAAAAAATAATTTATCATATTTTCA 4000

QY 964 AATTAATCACTCAATATTTCCAGTGAATAGCAAGAAATTTTACATATTAATCCAT 1023
DB 3999 TTTAAACATTTTCAATTTTTTTTTTTTTTAAACATTAATTAATTTTAAATTTTAA 3940
QY 1024 TTTTATCTCTGCTGCTCAAGTTTGTGTAGAATAGCAAGATCTGCTGATAT 1083
DB 3999 AATTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAA 3887
QY 1084 AGTACTGCTGGTGAATTAAGAACTGTGTAATCATTTGAAACAAATTAAGTAATGA 1143
DB 3886 ATTTCAATTTTTTTATCATTAACAAATCAATTTTCAATTTTCAATTTTCAAA 3827
QY 1144 TGTCTTAAATTAATTAATACATCAAGATATTAAGTGTACAGTTGAGCTGATTA 1203
DB 3826 AAAACAAAAATTCACAAATTAATTAATTAATTTTAAATTAATTTTCAATTTCAAA 3767
QY 1204 GAACATTAATGCAATTTGATGGAATATCGCTGAAAAATTTAGTAAATATGCTGT 1263
DB 3766 AAAATTAATTTTAAATAATTAATTAATTTTATTTTATTTTATTAATTAATTTAA 3707
QY 1264 TGCATCTGTTAATCTAAAGCATGAAATTAACCTATTAATTAATTAATTAATTAATCA 1323
DB 3706 TCATTTTAAATAATTAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTA 3647
QY 1324 TTCAATAGATCTTCATACATTAATTTGCTCATCATGCTTGGAGATCTTAATTAATCATG 1383
DB 3646 TACATTAATTTTCAATTAATTAATTTTATTAATTTTAAATTAATTTTATTTTATTAAT 3587
QY 1384 TACACAGAACTTCAGCAATATGCTTTTGAATTAATTAATTAATTAATTAATTAAT 1443
DB 3586 TCATTTTAAATAATTTTCAACAAATTAATTTTAACTTAATTTTATTAATTAATTAAT 3527
QY 1444 AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1503
DB 3526 AATTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3475
QY 1504 GATTAATTTGCAAAAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTT 1563
DB 3474 ATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTT 3415
QY 1564 ATTAACAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTT 1623
DB 3414 AAAAAAATTAATTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 3355
QY 1624 CTTTGAATTTTCAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTT 1683
DB 3354 ATTCATTTTATTAATTAATTAATTAATTTTATTTTATTTTATTTTATTTTATTTTATTT 3295
QY 1684 CCACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTT 1743
DB 3294 ACATAAATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTT 3235
QY 1744 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTCT 1803
DB 3234 TAAATTAATTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTAT 3176
QY 1804 ATCTCTATCTCAATTAATTAATTTGATTAATTAATTAATTAATTAATTAATTTTAA 1863
DB 3175 ATTTAAATTAATTAATTAATTAATTTTATTAATTAATTAATTAATTAATTAATTTAA 3116
QY 1864 CTTATTTTAAATTAATTAATTTTGAACATTTTCACTCATCATCATTAATTAATTAATTA 1923
DB 3115 ATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTAA 3056
QY 1924 AATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTAC 1983
DB 3055 AATTTATTTATTTTATTAATTAATTTTATTAATTAATTAATTAATTAATTAATTTAT 2996
QY 1984 TATTTACTAGTAATTTATCATATGATGCTGACATTAATTAATTTATCATCTTGACAT 2043
DB 2995 AATTTTCAATTAATTTTATTTTATTTTATTTTATTAATTAATTAATTTATTTTATTT 2936

Db	818	CATTATTAATTTATTTTAAATATATATTAATTAATTTTATTAATTTTAAATTT	759
Qy	4247	ACGCCAGACGCAATTAAGATTCATATTATGGTAATTTTATTTTAAACACCTAGCAAAAT	43066
Db	758	TAAATATTTAAAAAATTTTAAATTTAAATATTAATTTTAAATTTTAAATTTTAAAAA	699
Qy	4307	AATATCCCAACCAATGACTATCAACATATGTGATGATGATAAACAATCCACCAACAGAGT	43566
Db	698	AATTTTTTATATTTATTTATTTTATTTTATTTTATTTTAAATTAATTTTAAAAATTTT	639
Qy	4367	ATTATCATTTGACTACTTCACGTAATAAATCTAGATGAGTTAACGACATCTCCATGTAT	44287
Db	638	TTTAAAAAATTTTAAATTTTATATATATTTT---TTTATTAATTAATTTTATTTTAAATATA	582
Qy	4427	TGAAGTCAATCATCCGCTATATTATTCATTTATTCACATTAATATGTCCAAAACATCAAT	44868
Db	581	AATATATTTTAAATTAACATTTTATATTAATTTATATTAATAAATTTTAAATTTTAAATATAT	522
Qy	4487	GATGAATTTTGTATCATGATACCAATTTTGCAAAACAGAAATTTCTTGACTGTAT	45468
Db	521	TTAATATATTAATAATTTTAAATTTTATTTTATTTTAAATTAATTTTATTTTAA	462
Qy	4547	CTGTAAAGTATTTGTGTAAACAGCCTTTTACTGAGCAATTTTGATCATGATGACTAGA	46066
Db	461	AATTAATTTTATTTTATTTTATTTTAAAAAATAAATAAATTAATTAATTAATTAATTA	402
Qy	4607	CTCATTTTGTTCCAAT	4626
Db	401	TTAAATTTTAAATTTAAAT	382

```
RESULT 8
ACA29483
ID ACA29483 standard; DNA; 1029 BP
```

DT	19-JUN-2003	(first entry)
XX		
DE	Prokaryotic essential gene #111140.	

KW Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.

OS Clostridium difficile.

PN WO200277183-A2.
YY

03-OCT-2002
PD
XX

21-MAR-2002; 2002WC-050091U/

PR 06-SEP-2001; 2001US-00948993.

PR 08-FEB-2002; 2002US-00072851

2 X
3 X
4 X
5 X
6 X
7 X
8 X
9 X
10 X
11 X
12 X
13 X
14 X
15 X
16 X
17 X
18 X
19 X
20 X
21 X
22 X
23 X
24 X
25 X
26 X
27 X
28 X
29 X
30 X
31 X
32 X
33 X
34 X
35 X
36 X
37 X
38 X
39 X
40 X
41 X
42 X
43 X
44 X
45 X
46 X
47 X
48 X
49 X
50 X
51 X
52 X
53 X
54 X
55 X
56 X
57 X
58 X
59 X
60 X
61 X
62 X
63 X
64 X
65 X
66 X
67 X
68 X
69 X
70 X
71 X
72 X
73 X
74 X
75 X
76 X
77 X
78 X
79 X
80 X
81 X
82 X
83 X
84 X
85 X
86 X
87 X
88 X
89 X
90 X
91 X
92 X
93 X
94 X
95 X
96 X
97 X
98 X
99 X
100 X
101 X
102 X
103 X
104 X
105 X
106 X
107 X
108 X
109 X
110 X
111 X
112 X
113 X
114 X
115 X
116 X
117 X
118 X
119 X
120 X
121 X
122 X
123 X
124 X
125 X
126 X
127 X
128 X
129 X
130 X
131 X
132 X
133 X
134 X
135 X
136 X
137 X
138 X
139 X
140 X
141 X
142 X
143 X
144 X
145 X
146 X
147 X
148 X
149 X
150 X
151 X
152 X
153 X
154 X
155 X
156 X
157 X
158 X
159 X
160 X
161 X
162 X
163 X
164 X
165 X
166 X
167 X
168 X
169 X
170 X
171 X
172 X
173 X
174 X
175 X
176 X
177 X
178 X
179 X
180 X
181 X
182 X
183 X
184 X
185 X
186 X
187 X
188 X
189 X
190 X
191 X
192 X
193 X
194 X
195 X
196 X
197 X
198 X
199 X
200 X
201 X
202 X
203 X
204 X
205 X
206 X
207 X
208 X
209 X
210 X
211 X
212 X
213 X
214 X
215 X
216 X
217 X
218 X
219 X
220 X
221 X
222 X
223 X
224 X
225 X
226 X
227 X
228 X
229 X
230 X
231 X
232 X
233 X
234 X
235 X
236 X
237 X
238 X
239 X
240 X
241 X
242 X
243 X
244 X
245 X
246 X
247 X
248 X
249 X
250 X
251 X
252 X
253 X
254 X
255 X
256 X
257 X
258 X
259 X
260 X
261 X
262 X
263 X
264 X
265 X
266 X
267 X
268 X
269 X
270 X
271 X
272 X
273 X
274 X
275 X
276 X
277 X
278 X
279 X
280 X
281 X
282 X
283 X
284 X
285 X
286 X
287 X
288 X
289 X
290 X
291 X
292 X
293 X
294 X
295 X
296 X
297 X
298 X
299 X
300 X
301 X
302 X
303 X
304 X
305 X
306 X
307 X
308 X
309 X
310 X
311 X
312 X
313 X
314 X
315 X
316 X
317 X
318 X
319 X
320 X
321 X
322 X
323 X
324 X
325 X
326 X
327 X
328 X
329 X
330 X
331 X
332 X
333 X
334 X
335 X
336 X
337 X
338 X
339 X
340 X
341 X
342 X
343 X
344 X
345 X
346 X
347 X
348 X
349 X
350 X
351 X
352 X
353 X
354 X
355 X
356 X
357 X
358 X
359 X
360 X
361 X
362 X
363 X
364 X
365 X
366 X
367 X
368 X
369 X
370 X
371 X
372 X
373 X
374 X
375 X
376 X
377 X
378 X
379 X
380 X
381 X
382 X
383 X
384 X
385 X
386 X
387 X
388 X
389 X
390 X
391 X
392 X
393 X
394 X
395 X
396 X
397 X
398 X
399 X
400 X
401 X
402 X
403 X
404 X
405 X
406 X
407 X
408 X
409 X
410 X
411 X
412 X
413 X
414 X
415 X
416 X
417 X
418 X
419 X
420 X
421 X
422 X
423 X
424 X
425 X
426 X
427 X
428 X
429 X
430 X
431 X
432 X
433 X
434 X
435 X
436 X
437 X
438 X
439 X
440 X
441 X
442 X
443 X
444 X
445 X
446 X
447 X
448 X
449 X
450 X
451 X
452 X
453 X
454 X
455 X
456 X
457 X
458 X
459 X
460 X
461 X
462 X
463 X
464 X
465 X
466 X
467 X
468 X
469 X
470 X
471 X
472 X
473 X
474 X
475 X
476 X
477 X
478 X
479 X
480 X
481 X
482 X
483 X
484 X
485 X
486 X
487 X
488 X
489 X
490 X
491 X
492 X
493 X
494 X
495 X
496 X
497 X
498 X
499 X
500 X
501 X
502 X
503 X
504 X
505 X
506 X
507 X
508 X
509 X
510 X
511 X
512 X
513 X
514 X
515 X
516 X
517 X
518 X
519 X
520 X
521 X
522 X
523 X
524 X
525 X
526 X
527 X
528 X
529 X
530 X
531 X
532 X
533 X
534 X
535 X
536 X
537 X
538 X
539 X
540 X
541 X
542 X
543 X
544 X
545 X
546 X
547 X
548 X
549 X
550 X
551 X
552 X
553 X
554 X
555 X
556 X
557 X
558 X
559 X
560 X
561 X
562 X
563 X
564 X
565 X
566 X
567 X
568 X
569 X
570 X
571 X
572 X
573 X
574 X
575 X
576 X
577 X
578 X
579 X
580 X
581 X
582 X
583 X
584 X
585 X
586 X
587 X
588 X
589 X
590 X
591 X
592 X
593 X
594 X
595 X
596 X
597 X
598 X
599 X
600 X
601 X
602 X
603 X
604 X
605 X
606 X
607 X
608 X
609 X
610 X
611 X
612 X
613 X
614 X
615 X
616 X
617 X
618 X
619 X
620 X
621 X
622 X
623 X
624 X
625 X
626 X
627 X
628 X
629 X
630 X
631 X
632 X
633 X
634 X
635 X
636 X
637 X
638 X
639 X
640 X
641 X
642 X
643 X
644 X
645 X
646 X
647 X
648 X
649 X
650 X
651 X
652 X
653 X
654 X
655 X
656 X
657 X
658 X
659 X
660 X
661 X
662 X
663 X
664 X
665 X
666 X
667 X
668 X
669 X
670 X
671 X
672 X
673 X
674 X
675 X
676 X
677 X
678 X
679 X
680 X
681 X
682 X
683 X
684 X
685 X
686 X
687 X
688 X
689 X
690 X
691 X
692 X
693 X
694 X
695 X
696 X
697 X
698 X
699 X
700 X
701 X

XX	Wanda T.	Zamudio C	Mat
PT			

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX

DR WP1; 2003-029926/02
DR P-PSDB; ABU25613.

PT New antisense nucleic acids, useful for identifying proteins or screening

PT isolate candidate molecules for rational drug discovery programs

PS Claim 14; SEQ ID NO 17353; 1766pp; English.

CC The invention relates to an isolated nucleic acid comprising any one of

the 621 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-regulated gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the targete prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WFO at [ftp.wipo.int/pubd/published_pat_sequences](http://wipo.int/pubd/published_pat_sequences)

Query Match	Score	DB	Length
3.7%	175.2	7	1029

Matches 483; Conservative 0; Mismatches 438; Indels 15; Gaps 3

```

Qy      634 GAATACAGATTTTAGGTCAGACTCATTTGGCCAGCATAAGCAATTGCCTGTACG 693
        |||||       |   |   |   |   |   |   |   |   |   |
Db      18 GAAAGTATGTGTAATTAGGTACTGGAAGCTGGGAAGTGCAATTAGCGCTTGCAA 77

```

694 ACATGGTATATCAGTTAACTTATGGGACGTGATCATAGAATAATTACACATATAACAC 753

Dd 78 A A A G C T A T G A T G T G A G C A T G T G G A C T C G A A A G A G G A C A G C T A A A A A A A I A A C A G 13

754 T T A C C G A A A A A A I I T A A A A I A I I T A C C C A C A I A I C A I C A C C G A C A C A C A I A I A I A G C A A C 813

22 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523</
--	---	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-------

[illegible]

308

934 TCCGATATTTAGTAAAGGTATCGAAATTACATCACTCAAAATTTCCCGAGTGAAT 993

309 AGTTATTGTTAATGTAGCTTAAAGGCTTTGAAAAGGACTGGGCTTAGATTATCACAAGT 368

994 AGCAGAAGAAATTTACAATAATCCAATTTTATACTCTCTGGTCCAGTTTGCTAA 105

Db 369 CTGTGAGAGAAATTACCACAAATCCATATGTAATACTATCTGGCCATCTCATGCTGA 428

1054 AGAATTGCAGACATCTTCCTGTAGTATAGTACTTGCTGGTGATATAAGAACTTGG 111

Db 429 AGAAGTAGCAAGGATATTCCTACTACCGTAGTGTGCTTCTAAAGATTAAAAATAGC 488

1114 TGAATCATGTAGAAACAATAAGTAATGATGTTCTAAAAATATATACCATCAAGATAT 1115

D5 489 ACAGATGATACAAGATTTGTTTATGAGCCCTAAACTTAGAGTTTATACGATTCGAGCA 578

QY 1174 TATAGGTGACAGATTGGAGCTGCTAATAAGAACATAATTGCAATTGCGATGNGAATAT 1233
 DB 549 AGTGGAGTAGAAGCTGGAGGAGCATTTAAAAATATATAGCATTTGGAGCTGGGATATG 608
 QY 1234 CCGTGAAGAAAAATTTAGGTAATTAATGCTGCTCTACTGTATTAAGTAAAGCATGATGA 1233
 DB 609 TACGAGATTAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 668
 QY 1294 AATTAACACATATATATAGCAAAAAAATCATTCATTCATTCATTCATTCATTCATTCAT 1353
 DB 669 AATGAGACACATTCGATTCG-----TATGGAGCTAATATGCTTACCTTGGAGGCT 722
 QY 1354 ATCAGTCTTGAGATCTAATTTAATCAATGTCACAGAACCTTACCGCATATGCGCTT 1413
 DB 723 TTAGGATATGAGATTTATATGTTACTGTACAAAGTATGCTAGTGAAGATAGAGAGC 782
 QY 1414 TGGACTAGAAATAGCAAAAAAGTATGAAATTAATATCATTTAATGATCACAACCTTAAAGCT 1473
 DB 783 TGGTACTCATATAGGTAAGGATGAGTATGAGATATCTTTAAAGAGATGAAAGATGCT 842
 QY 1474 TGTGAGAGACACAGTACTGTAAACCACTGATTCATTCATTCATTCATTCATTCATTCAT 1533
 DB 843 TGTAGAGAGACATACCTGCTACTGAAAGTACACATGATGCTGTGAAATTAAGATATGA 902
 QY 1534 ACTACCAATTTGCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1569
 DB 903 TATGCTATTAATTAATGCTATATATTCCTATATATA 938

RESULT 9
 ID ACA47630
 ACACA7630 standard; DNA; 996 BP.
 ACACA7630;
 19-JUN-2003 (first entry)
 DE Prokaryotic essential gene #29287.
 XX
 KM Antisense; ds: prokaryotic essential gene; cell proliferation;
 XX drug design; gene.
 OS Stephylococcus haemolyticus.
 XX
 XX W020027183-A2.
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002MO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-0007851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH,
 XX WPI; 2003-02926/02.
 DR P-PDB; ABU43760.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 14; SEQ ID NO 35500; 1766pp; English.
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:

CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which the test compound required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
 CC prokaryotic essential genes. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 996 BP; 366 A; 119 C; 202 G; 309 T; 0 U; 0 Other;
 XX
 SQ
 Query Match 3.6%; Score 170.6; DB 7; Length 996;
 Best Local Similarity 49.6%; Pred. No. 4.9e-18;
 Matches 468; Conservative 0; Mismatches 469; Indels 6; Gaps 1;

QY 635 AAAAAATGATTTTATGCTGAGATGATTTGGACAGCATATGCAATTTGACGCTGATGA 694
 DB 7 AAAAAATGATTTTATGCTGAGATGATTTGGACAGCATATGCAATTTGATGATGATGA 66
 QY 695 CATGATATATGATTTTATGCTGAGATGATTTGGACAGCATATGCAATTTGACGCTGATGA 754
 DB 67 AATGACATATCTTTTATGCTGAGATGATTTGGACAGCATATGCAATTTGATGATGAT 126
 QY 755 TACGAAAAATTTTAAATATTTTACCACATATCTACACAGCATATGCAATTTGATGAT 814
 DB 127 CACCATCAAAATTAAGATTTTAAAGATGCTGATTTGATGATGATTTAAAGGAACT 186
 QY 815 AGCAATATAGCAGATTTATCTGACACAAATATGATTTTAACTATTTCTTACATTTCTTACA 874
 DB 187 AGTATTTTAAAGACCGGAAATTTTACGATATATTTTAAAGCTTACCAACTTAA 246
 QY 875 CAACAATTCAGCACCATATGATACAAATACAAACAAACAGCATATGCTTAAATTTACT 934
 DB 247 GCATGCGCGAAGTTCAAGTGAATTTGATTTTAACTTTTAACTTTTAACTTTTAACT 306
 QY 935 CCAATATTTATTTGATTTTAAAGTATGCAATTTTAACTTTTAACTTTTAACTTTTAACT 994
 DB 307 CAGTTGCAAAAGATATGATTTTAACTTTTAACTTTTAACTTTTAACTTTTAACTTTTAA 366
 QY 995 GCAGAGAAATTTTAACTTTTAACTTTTAACTTTTAACTTTTAACTTTTAACTTTTAA 1054
 DB 367 TCAATTTCAAGATATGATTTTAACTTTTAACTTTTAACTTTTAACTTTTAACTTTTAA 426
 QY 1055 GAAATTCAGACATCTCTTGTATGATTTTAACTTTTAACTTTTAACTTTTAACTTTTAA 1114
 DB 427 GAGGATGATTTTAACTTTTAACTTTTAACTTTTAACTTTTAACTTTTAACTTTTAACT 486
 QY 1115 GAATCATTTAGTAAATTAATTTTAACTTTTAACTTTTAACTTTTAACTTTTAACTTTTAA 1174
 DB 487 AAATGATTTCAAGATTTTAACTTTTAACTTTTAACTTTTAACTTTTAACTTTTAACT 546
 QY 1175 ATAGGCTAGATTTGAGATTTTAACTTTTAACTTTTAACTTTTAACTTTTAACTTTTAACT 1234

Db 547 GTGCGTTGAACTAGAGAGACATTAAAAATATATATTGCACTTGCAAGTGTATTGTA 606
 Qy 1235 GCTGGAATAAATTAGTAATATATGCTGTGTCTACTGTATTAATAAGCATGATGAA 1294
 Db 607 GCAGGATGAGGCTATGATATGCTAAGCTGTTATGACAGAGTTAGCTGAA 666
 Qy 1295 ATTAAACACTATATATAGCAAAAAATCATTCATATGATCTTCAATTAATTGTGCA 1354
 Db 667 ATAGCCGATTAAGTATAAAAATTAAGTGTCT-----GATCAATGACATTTTATGCTTT 720
 Qy 1355 TCATGCTTGTGAGATCTAATATTAATACATGTACAGACAAACATTCACGCAATATGCTTT 1414
 Db 721 GGTGATATTGAGATCTAATTTGTTACTGTACATCAAGCATTCACGTAATATACCTTTA 780
 Qy 1415 GGACTGAATATAGAAAGAGTAAAAATATATATACCTTATATGATCAACACTAAAGCTT 1474
 Db 781 GGAATTAATTAAGGACAAAGCCAAACAAATGATGAAGCTTTAAATGAATGAAATGGA 840
 Qy 1475 GTTGAAGAACAGTACTGTAAAAACCACTGATATCATTTAGCAAAAAAATTATGTAGAA 1534
 Db 841 GTAGAGGAGATTATATCTATCTATCTGTATACATTTAGCTAAACAAATATGAT 900
 Qy 1535 CTACCAATTTGCAATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCT 1577
 Db 901 ATGCTTATATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCT 943

RESULT 10

ACF75135
ID ACF75135 standard; DNA; 996 BP.

ACF75135;

DT 20-NOV-2003 (first entry)

DE Staphylococcus aureus DNA #2815.

KM Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;

XX enzymatic assay; antibiotic target; gene; ds.

OS Staphylococcus aureus.

PN WO200294868-A2.

PD 28-NOV-2002.

PF 27-MAR-2002; 2002WO-IB002637.

PR 27-MAR-2001; 2001GB-00007661.

PA (CHIR-) CHIRON SPA.

PI Masignani V, Mora M, Scarselli M;

DR WPI; 2003-120786/11.

PT P-PSDB; ABM73575.

PT New Staphylococcus aureus protein, useful as a vaccine for treating or

PT preventing Staphylococcal infection, specifically an infection caused by

PT S. aureus, e.g. sepsis.

PS Claim 6; SEQ ID NO 5629; 49pp; English.

CC The invention relates to novel genes and encoded proteins from
 CC Staphylococcus aureus. A composition comprising the S. aureus protein, a
 CC nucleic acid encoding the protein, or an antibody to the protein, is
 CC useful as a pharmaceutical, particularly as a vaccine for treating or
 CC preventing infection due to Staphylococcus bacteria, specifically an
 CC infection caused by S. aureus. The composition is particularly useful for
 CC treating or preventing sepsis in a patient. The composition can also be
 CC used for diagnostics. The protein is also used in an assay for enzymatic
 CC studies and as a target for antibiotics. This sequence represents one of
 CC the novel S. aureus genes of the invention

XX Sequence 996 BP; 345 A; 134 C; 210 G; 307 T; 0 U; 0 Other;
 SO Query Match 3.5%; Score 168.6; DB 7; Length 996;
 Best Local Similarity 51.4%; Pred. No. 1e-17;
 Matches 505; Conservative 0; Mismatches 454; Indels 24; Gaps 4;

Qy 635 AAAATCGTATTTTATAGTGAAGATCATTTGGACAGCAATAGCAATGATGATGATGACGA 634
 Db 7 AAAATTCGCTTTTGTATGAGAGTTTGGACAGCCCTTGCCAAATGTTCTTGCAAA 66
 Qy 695 CANGATATCAGTAACTTATGAGGAGCGATCATAGAAATTTACATATATAACACT 754
 Db 67 AATGACATGATGTTTATGATGAGTGGTAAATCAAGATGCTGTATGAAATTAATGACA 126
 Qy 755 TACGAAAAATTTAAATTTTAAACCATATCATCTACAGACAAATATGCAAC 814
 Db 127 TGTCTATCAAAATTAAGATTTTAAATACGCAAAATTAAGTGTAAACATCATGCTACT 186
 Qy 815 AGCAATATAGACGAGATATATCTGACAAACATACATGATTTATCTTAATCTCTACA 874
 Db 187 TCGATATGACAAAGCATCTCAATTTGCGATATTTACTTAATG---CTTACCTACT 243
 Qy 875 CAACAATTAAGCAACATATGATACAAATTAACAACAGCAATATGTTAAATTAAT 934
 Db 244 AAAGCAATGAGAGAGTGGCTCTCAAAATTAATGATTAAGC-----TGACCTTAATAAG 297
 Qy 935 CCAATATTAATTTGTATAGGATATCGAAATTAATCACTCACTCAATTTCCAGTGAATA 994
 Db 298 ACTTTTATCATGTTGCTAAGGATATTAAGGAAATGGAACGTTTAAAGTGTGAGAAATG 357
 Qy 995 GCAGAGAAATTTT-----ACAATATATCCAAATTTTATCTCTGCTGTCAGT 1045
 Db 358 ATGGAATATCTATTTCACTGATATTAATGCAAGTATTTGGCTGTGTCAGGCGCAAGT 417
 Qy 1046 TTGCTTAAGAAATTTGAGAAACATCTTCTGTATGATGATGCTGCTGATTAATTA 1105
 Db 418 CATGCGAAGAAATTTGATGCAACCAACCACTACAGTCTCTCATCAAAAGATTA 477
 Qy 1106 GAACCTGTGATCATGATGAGAAACAAATGATGATGATGTTTAAATTAATTAATCAAT 1165
 Db 478 AGTGTAGTAAATTAAGCAAGATTATTAATTAATTAATTAATTAATTAATTAATTAAT 537
 Qy 1166 CAAGATTTATAGTGTACAGATTGAGCGCTTAAGAAACATTAATGCAATTAATGCAATG 1225
 Db 538 GATGACTGTATGATGTTGAACCTTGATGATGATGATGATGATGATGATGATGATGAT 597
 Qy 1226 GGAATTAATCGTGAATAAATTTAGTAAATATGCTGTTGCTAATGATTAATTAATTAAGC 1285
 Db 598 GGTATCTAGCTGAATATGCTACGGGTATGATGATGATGATGATGATGATGATGATGATG 657
 Qy 1286 ATGATGAATTAATAACATATATATAGCAAAAAATCATTCATATGATGATGATGATGAT 1345
 Db 658 TTAGCGGAAATTTAGTGA-----TTAGTGAAGATTTAGTGGCCATCTCATGACATTT 711
 Qy 1346 ATTGTCATCATGCTTTGAGAGATCTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1405
 Db 712 CTAGCTTTAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 771
 Qy 1406 ATGCTTTTGAATGAAATAGGAAAGATGAAATTAATTAATTAATTAATTAATTAATTAAT 1465
 Db 772 TTCACTTAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 831
 Qy 1466 CTAAAGCTTTGAGAGACAGTATGTAATAACCTGATATCATTTAGCAAAAAAATCTT 1525
 Db 832 AATATGTTGTGTAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 891
 Qy 1526 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1585
 Db 892 AATGTGATATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 951
 Qy 1586 GATTAAGCCATATCAACATATTT 1608

XX DE Prokaryotic essential gene #9182.
XX KM Antisense; ds: prokaryotic essential gene; cell proliferation;
XX KM drug design; gene.
XX OS Clostridium acetobutylicum.
XX PN WO200277183-A2.
XX PD 03-OCT-2002.
XX PF 21-MAR-2002; 2002WO-US009107.
XX PR 21-MAR-2001; 2001US-00815242.
XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0342923P.
XX PR 08-FEB-2002; 2002US-00072851.
XX PR 06-MAR-2002; 2002US-0362639P.
XX PA (ELIT-) ELITRA PHARM INC.
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KJ, Zyekind JW,
XX PI Mail D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX DR WPI: 2003-0289926/02.
XX DR P-PSDB; AB023655.
XX PT New antisense nucleic acids, useful for identifying proteins or screening
XX PT for homologous nucleic acids required for cellular proliferation to
XX PT isolate candidate molecules for rational drug discovery programs.
XX PS Claim 14; SEQ ID NO 15395; 1766pp; English.
XX XX The invention relates to an isolated nucleic acid comprising any one of
XX CC the 6213 antisense sequences given in the specification where expression
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:
XX CC (1) a vector comprising a promoter operably linked to the nucleic acid
XX CC encoding a polypeptide whose expression is inhibited by the antisense
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX CC polypeptide or its fragment whose expression is inhibited by the
XX CC antisense nucleic acid; (4) an antibody capable of specifically binding
XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX CC proliferation or the activity of a gene in an operon required for
XX CC proliferation; (7) identifying a compound that influences the activity of
XX CC the gene product or that has an activity against a biological pathway
XX CC required for proliferation, or that inhibits cellular proliferation; (8)
XX CC identifying a gene required for cellular proliferation or the biological
XX CC pathway in which a proliferation-required gene or its gene product lies
XX CC or a gene on which the test compound that inhibits proliferation of an
XX CC organism's activity; (9) manufacturing an antibiotic; (10) profiling a
XX CC compound's activity; (11) a culture comprising strains in which the gene
XX CC product is overexpressed or underexpressed; (12) determining the extent
XX CC to which each of the strains is present in a culture or collection of
XX CC strains; or (13) identifying the target of a compound that inhibits the
XX CC proliferation of an organism. The antisense nucleic acids are useful for
XX CC identifying proteins or screening for homologous nucleic acids required
XX CC for cellular proliferation to isolate candidate molecules for rational
XX CC drug discovery programs, or for screening homologous nucleic acids
XX CC required for proliferation in cells other than S. aureus, S. typhimurium,
XX CC K. pneumoniae or P. aeruginosa. The present sequence data for this target
XX CC prokaryotic essential genes. Note: The sequence data for this patent did
XX CC not form part of the printed specification, but was obtained in
XX CC electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 999 BP; 366 A; 123 C; 236 G; 274 T; 0 U; 0 Other;

Query Match 3.5%; Score 168.2; DB 7; Length 999;
Best Local Similarity 51.0%; Pred. No. 1.2e-17;
Matches 482; Conservative 0; Mismatches 448; Indels 15; Gaps 3;
648 TAGGTGAGATCATTTGGACAGCAATAGCAATTGCTGTCAGACATGTATATCAG 707

Db 20 TAGGAGAGAGAGCTTTGTACAGACATTCGCAATATGCTGTAAGAAAGGCATATG 79
Qy 708 TTAACCTTGGGACGTCATCATAGAAATATTAACATATTAACCTGACGAAATAT 767
Db 80 TTGTGATTTGGATAGAAATTAAGAAATTTCTGAAGATATTAATCAATTAAGACTATA 139
Qy 768 TAAATATTTTACCAATATCATCTACAGCAATATATATGCAACGACATATATAGCG 827
Db 140 CAAGATATCTTCCAAACAAATATATACCTGCTGCTTAAGACGATGATATATGGA 199
Qy 828 AAGTATTTCTGACAAATATCATGATATATCTTAATCTTCCATGACAAATTAAGCA 887
Db 200 AGG---CGCTAAGAAAGCAAAATATATATTTCTTGTGCTCCATCTTTGCTATTAAGGG 256
Qy 888 CCATATGTACAAATATACAAACAAACAGCATATGTGTAAATATCTCAATATTAATTT 947
Db 257 AAGTATGAGGAAATTAAGGGCTT-----TTAAGAAAGACCAATATATTAAGTA 310
Qy 948 GTAGTAAAGTATGAAATTAATCATCACTCAATTTCCAGTGAATATGCAAGAAATTT 1007
Db 311 TTGCAAAAGCATGGAAGAGAGAGACCAAGAGAGCTGTCTGAAGTATTAAGAGAGAC 370
Qy 1008 TACATATATATTCATTTTATATCTCTGCTGCTCAAGTTTGTAAAGAAATGACAGAC 1067
Db 371 TTTATAAATATCTGTTGTATGATCTTTCTGCTCTTCTGCTGATGCAAGAAATG 430
Qy 1068 ATCTTCTTGTATGATGATGATCTGCTGCTGATTAATTAAGAACTTGATATG 1127
Db 431 ATATCCACATACAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 490
Qy 1128 AAACATAGTAAATGATGCTTAAATTAATATATCAATCAATATATATAGGTGACAG 1187
Db 491 ATGTTTTATATGCTACAGCTTTTATAGGTGTACAAATATGATATGAGCTTGAA 550
Qy 1188 TTGAGCTGCTATTAAGAAATTAATTTGCAATGATGATGATATATGCTGAAATAT 1247
Db 551 TAGGAGAGAGCTTAAATTAATTTATGCTGATGATGATGATGATGATGATGATG 610
Qy 1248 TAGGATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1307
Db 611 ATGTGACAAACAAAGCGCGCTTATGACAAAGGATATGATGATATATGAGA---- 666
Qy 1308 ATATGCAAAATATCATCATATGATGATGATGATGATGATGATGATGATGATG 1367
Db 667 --ATAGGGGTAAAGGTGAGAGTAAAGCAAACTTTTGGACTTACAGAGCTGAGTG 724
Qy 1368 ATCTATATTTAATCATGATGATGATGATGATGATGATGATGATGATGATGATG 1427
Db 725 ACCATATGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 784
Qy 1428 GAAAGGTAGAAATTAATTAATCATTAATGATGATGATGATGATGATGATGATG 1487
Db 785 GTAGGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 844
Qy 1488 GTACTGTAAACCACTGATATCATTTAGCAAAATTAATGATGATGATGATGATG 1547
Db 845 AAGCGTGTACACCTTTATGAAATTAAGAAATCTTAGGGTATCTATGTCATATACCA 904
Qy 1548 TATCTATTTAATTAATTTATTAATGATGATGATGATGATGATGATGATGATG 1592
Db 905 CTTCACCTTATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 949

RESULT 13
AAX20248_00/c
WP Sequence split into 10 fragments LOCUS AAX20248 Accession Aax20248
WP Fragment Name Begin End
WP AAX20248_00 1 11000
WP AAX20248_01 100001 210000
WP AAX20248_02 200001 310000
WP AAX20248_03 300001 410000
WP AAX20248_04 400001 510000

Continuation (2 of 10) of AAX20248 from base 100001 (Borrelia burgdorferi polynucleotide
 WP Sequence Split Into 10 Fragments LOCUS AAX20248 Accession AAX20248

Fragment Name	Begin	End
WP AAX20248_00	1	110000
WP AAX20248_01	100001	210000
WP AAX20248_02	200001	310000
WP AAX20248_03	300001	410000
WP AAX20248_04	400001	510000
WP AAX20248_05	500001	610000
WP AAX20248_06	600001	710000
WP AAX20248_07	700001	810000
WP AAX20248_08	800001	910000
WP AAX20248_09	900001	910715

Query Match 3.5%; Score 165.8; DB 2; Length 110000;
 Best Local Similarity 49.4%; Pred. No. 3,4e-17;
 Matches 598; Conservative 0; Mismatches 577; Indels 36; Gaps 5;

QY	DB	Fragment Name	Begin	End
QY 2661	DB 6138	TAATAGGCGAATAGACTATATCTAGTGTAACTTTGGCTGTATCATCTCTTAC	2720	2720
QY 2721	DB 6081	GAGATTCATTAATTAATCCATCTTAACATTACGTAATGCTTAAAGTTACCTGGCC	2780	2780
QY 2781	DB 6021	ATTCATATGCTGATGCTATTAAGCTTCACTTAATCATAGAGACAAACCTA	2840	2840
QY 2841	DB 5961	TTTTTTACAGATGCTATTCATTAATTAACCTAACAATCCGATATCTGTACAATAT	2900	2900
QY 2901	DB 5901	CTACTAAGCGGTACTCTTAATTTGAGAGACATTTAATTAATTAATTTCTCAACA	2960	2960
QY 2961	DB 5841	ATTTTCCTTTTAATTTCTCTCAATGTTTGTGTTGTCAGACGAGACCTGATAT	5782	5782
QY 3021	DB 5781	CAACTTAATTTAGTCTCTCCCAACAGCTTCAAAAGTTTGTGTCAGTACTCTTA	5722	5722
QY 3081	DB 5721	AAAGCTTGACTTAATTTGAGAGATTTCTGCTATCTCAACGAAATTTGCTT	5662	5662
QY 3141	DB 5661	TGTTTCAAGTCAATCTGCTTTTGTGAAATGCTCCAGTATGCT-----C	5608	5608
QY 3201	DB 5607	CTTTTCAAGGCGAATTTGCTTCAAGATGCTTTCAAGAAAGTGTGCGCAATTTA	5548	5548
QY 3261	DB 5547	CTTTTAAATGTTTGTGCTTTTCTTAATTTGAAAGATCAAAAATACATCTGCTATTTT	5488	5488
QY 3321	DB 5487	CTTTTCAACAGCTGCGAAGCAGTATGATAGATGATGATGATGATGATGATGATGAT	5428	5428
QY 3381	DB 5427	TATCACTTCTCAATTAAGTATACAGGAGGAGTTACGATATTTCAATATCTTCAA	3440	3440
QY 3441	DB 5367	TTAGATCTTTTCTTAATTTTCTTAATTTGCAATTTTCAATTTTCAATTTTCTT	5308	5308

QY	DB	Fragment Name	Begin	End
QY 3484	DB 5307	TTTATTTAGGATCTTTTATTTATTTAGCAAAAGTCTTTCAAGGCTTAAGGCTTGTGTA	5248	5248
QY 3540	DB 5247	AAATATCAATGACACCTCTCTCAATGCACTACAGAGATCACTGTTCCGCGCTG	5188	5188
QY 3600	DB 5187	TTAGATATATAAAGATATCCCAAGTTTCTTTTCAATTTTGAAGCAATTTTCTC	5128	5128
QY 3660	DB 5127	CAGATATCTGGGCAATCT-----CAGGTCAGATATTTATTAATCAATCAATTTTCAATTTT	5074	5074
QY 3720	DB 5073	CAATTTTCAAGGCTTCTTCCGCTCACTAGAGAGAAACAAATATCTTCAATCT	5014	5014
QY 3780	DB 5013	CAAGATATAGCAATTTCTCTTAATTTCTTTTCAATCATCATCAAGATTTTGC	4954	4954
QY 3840	DB 4953	TCATTTTCAAT 4943		

RESULT 15
 ID ACA46288 standard; DNA; 999 BP.

ACA46288;

19-JUN-2003 (first entry)

Prokaryotic essential gene #27945.

Antisense; ds; prokaryotic essential gene; cell proliferation;

Staphylococcus aureus.

MO200277183-A2.

03-OCT-2002.

21-MAR-2002; 2002WC-US009107.

21-MAR-2001; 2001US-00815242.

06-SEP-2001; 2001US-00948993.

25-OCT-2001; 2001US-0342923P.

08-FEB-2002; 2002US-00072851.

06-MAR-2002; 2002US-0362699P.

(BLIT-) ELITRA PHARM INC.

Wang L, Zamudio C, Malone C, Haseelbeck R, Onsen KU, Zyskind JW;

Wang L, Trawick JD, Carr GJ, Yamamoto R, Foreyth RA, Xu HH;

WPI; 2003-029926/02.

P-FSDB; ABU42418.

New antisense nucleic acids, useful for identifying proteins or screening

for homologous nucleic acids required for cellular proliferation to

isolate candidate molecules for rational drug discovery programs.

Claim 14; SEQ ID NO 34158; 1766bp; English.

The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression

CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-regulated gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WFO at
 CC ftp.wipo.int/pub/published_pct_sequences
 CC XX

Sequence 999 BP; 349 A; 133 C; 210 G; 307 T; 0 U; 0 Other;

Query Match 3.5%; Score 165.4; DB 7; Length 999;
 Best Local Similarity 51.2%; Pred. No. 3.2e-17;
 Matches 503; Conservative 0; Mismatches 456; Indels 24; Gaps 4;

QY 635 AAAATCACTATTTAGTGTGAGATCTTGGACAGCAATGCAATTCGCTGCA 694
 DB 7 AAAATCACTATTTAGTGTGAGATCTTGGACAGCAATTCGCTGCAATTCGCTGCA 66
 QY 635 CATGTATATCACTATTTAGTGTGAGATCTTGGACAGCAATTCGCTGCA 754
 DB 67 AATGACATGATGTTGATGAGGGGTAAATCAAGTGTGATGAAATTAATCA 126
 QY 755 TACCGAAAAATTTAAATATTTTCCACATATCATCTACAGACACATATATGCAACC 814
 DB 127 TGTCTACAAATTAATAATTTTAAATACCGAAATTAAGTGTACATCATGCTACT 186
 QY 815 AGCAATATAGACGAATATATCTGACAAATATCATGATATCTTAATCTTCTACA 874
 DB 187 TCAGATATGACCAAGCAATTCATTCAGATATCTTAATGAG--CTTTACTACT 243
 QY 875 CAACATATACGACCATATGTAACAATATACAAACAGCATATGTGTAATAATCT 934
 DB 244 AAGCAATGCGAAGATGCTACTCAATTAATGATAGC-----TGACCTTAATAAG 297
 QY 935 CCAATATTAATTTAGTAAAGATTCGAAATTCATCATCAATTTCCAGTGAATA 994
 DB 298 ACTTTATACATGTTGCTAAAGATTTGAAATGAGAGCGTTAAACGTGTGCAAGATG 357
 QY 995 GCAGAGAAATTT-----ACAATATATCAATTTTAACTCTGCGGCAAGT 1045
 DB 358 ATTGAAGATCTATTTACCCGAATATATGACAGTATTCGCGTGTTCAGGCCAAGT 417
 QY 1046 TTGCTAAAGAAATTCGAAACATCTTCTGTATATAGTACTGCTGTGATATATA 1105
 DB 418 CATGGGAAAGATGTTGATGCAAGCAACATCAAGTGTGATCATCAAAAGATATA 477
 QY 1106 GAATCTGGGATCTATGATGAAACATATAGTATGATGTTCTAAATAATATATACAT 1165
 DB 478 AGTGAAGTAAATTAACGCAAGATTTATTTATGAATGATTTTTCGCTGTGTACAGAA 537
 QY 1166 CAAGATATATATAGTGTACAGATGAGCTGATTAAGAAACATATATGCAATTCATGT 1225

DB 538 GATGATTTAGTGTGTTGAACCTTGTTGATGTAAGAAATATCATGCACTAGCAAGT 597
 QY 1226 GGAATATATGCTGTAAGAAAAATTTAGTATATATGCTGTTGCTACTGTTTAACTAAAGC 1285
 DB 598 GGTATGTTAGCTGGAATTTGGCTACCGTGTATATGAAAAAGCTGCTTAATGATGCTGCTGT 657
 QY 1286 ATGATGAATATTAACACATATATATAGCAAAAAATCATCATATGATCTTCACTACTTA 1345
 DB 658 TTACCGAAATTAATAGTA-----TTAGTGAATAAGTTAGAGCCGATCTTATGCAATTT 711
 QY 1346 ATTGTCATCATGCTGTTGAGATCTTATATTTAACTATACATGTAACAGAACATTCAGCAAT 1405
 DB 712 CTAGGTTTATGTTGATGATCGTACTTAATCGTTACTTGTATATCAACATTCCTGAAAT 771
 QY 1406 ATGGCTTTTGAAGCTAAGAAATAGAAAAAGTAGAAATATTAATATATATATAGTACAAAC 1465
 DB 772 TTCACTATGATATATTAATTTGACAGAGTGAATCAATGATCAAGCATATCTGTAATG 831
 QY 1466 CTAAAGCTTTGTAAGAAACAGTACTGTAACCACTGATATCATTTAGCAAAAAAATCTT 1525
 DB 832 AATATGTTGTTGAAGATTTTATATCAATTAATGATTTATCATTTAGCTAAAGAAAA 891
 QY 1526 AATGTAACCTACCAATTTGCAATCTATTTTCAATTTATTCATGAGATATATACACTA 1585
 DB 892 AATGTAATATGCAATTTCAATTTCAATTTGATATATATGATATTTTGAATAATTTCTAGTA 951
 QY 1586 GATTAAGCCATATCAAAACATATT 1608
 DB 952 AAAGATGCTTAAGATTTAAT 974

RESULT 16
 ABZ10246
 ID ABZ10246 standard; DNA; 8056 BP.
 XX
 XX ABZ10246;
 XX
 DT 16-JAN-2003 (first entry)
 XX
 DE Haematopoietic cell proliferation disorder related DNA sequence #386.
 XX
 XX Human; haematopoietic cell proliferation disorder; cytostatic;
 KM Gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;
 KM cytosine methylation state; gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200277272-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 26-MAR-2002; 2002WO-EP003401.
 XX
 PR 26-MAR-2001; 2001US-0278333P.
 XX
 XX (EPIG-) EPIGENOMICS AG.
 XX
 PI Berlin K, Braun A, Dietler J, Guetig D, Howe A, Mueller J;
 PI Olex A, Piepenbrock C, Adorian P, Grabs G, Lesche R, Liu B;
 PI Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T, Perlet C;
 PI Schwope I, Ziebarth H;
 XX
 DR MPI, 2003-018942/01.
 XX
 XX
 PT Detecting and differentiating between hematopoietic cell proliferative
 PT disorders, comprises contacting a target nucleic acid with a reagent that
 PT distinguishes between methylated and non-methylated CpG dinucleotides.
 XX
 PS Claim 28; SEQ ID NO 386; 117bp; English.
 XX
 CC The present invention describes a method for detecting and
 CC differentiating between haematopoietic cell proliferative disorders

CC associated with at least 1 gene and/or their regulatory regions in a
 CC subject. The method comprises contacting a target nucleic acid in a
 CC biological sample obtained from the subject with at least 1 reagent,
 CC which distinguishes between methylated and non-methylated CpG
 CC dinucleotides within the target nucleic acid. AB209861 to AB211118
 CC represent specifically claimed nucleotide sequences from the present
 CC invention. Oligonucleotides from the present invention can be used: for
 CC differentiating between healthy haematopoietic cells and proliferative
 CC dysplastic haematopoietic cells; for differentiating between acute
 CC lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
 CC determining the cytosine methylation state and/or single nucleotide
 CC polymorphisms (SNPs) of haematopoietic cell proliferation disorder
 CC related sequences and their complements; and as primers for the
 CC amplification of haematopoietic cell proliferation disorder related DNA
 CC sequences. The nucleotide sequences from the present invention can also
 CC be used for detecting a predisposition to, differentiation between
 CC subclasses, diagnosis, prognosis, treatment and/or monitoring of
 CC haematopoietic cell proliferative disorders. The present method enables a
 CC highly specific classification of haematopoietic cell proliferative
 CC disorders allowing for improved and informed treatment of patients
 CC
 XX Sequence 8056 BP; 3711 A; 0 C; 371 G; 3974 T; 0 U; 0 Other;

Query Match 3.4%; Score 162.2; DB 7; Length 8056;
 Best Local Similarity 44.1%; Pred. No. 1.1e-16;
 Matches 1131; Conservative 0; Mismatches 1393; Indels 43; Gaps 9;

QY 172 ACCATTTCTCAATCTCCATTTTAACTGACAGAGTATTTATGTAAACAGACAC 231
 Db AAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 289
 QY 232 CCGTACTCTTATGACAGAACTCTATPAACCTTACAGGCTCTTACATAGAAACACT 291
 Db TTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 349
 QY 292 GATATCAAGATACCTCTCTGTCATGATATTTTCCATATTTTCAATATCA 351
 Db ATATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 409
 QY 352 TAGAGTCTTATGACATCATATTAATCAATCAAAACAAATTAAGTACTATCT 411
 Db TATATTTATTTATTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 463
 QY 412 ATTTATTTAAATTTCTCAATGCTTTTACAGAGACCTAAATTTATGACACTTA 471
 Db AAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 523
 QY 472 ACTTATTAATTTTCAATTTTATTTTACTACTCTCA--TATTTAGCATATATCA 529
 Db ATATTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 583
 QY 530 TCTGACATGTTTAACTAGTAAACATTTTCTTAACTAGATATATCAATTAAT 589
 Db TATTTATTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 643
 QY 590 ATATTTTCTGCTGTTCTGATATTAATTTAGTAAATTAATGAAATCGTATTTTA 649
 Db TTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 703
 QY 650 GGTGACGATCATTTGGACAGCATATGCACTGTCGACAGCATGATATCATGTT 709
 Db ATATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 763
 QY 710 AACTTATGGGACGATCATAGAAATTTATACATATTAACATTAACGAAATTAATTA 769
 Db TATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 823
 QY 770 AATTTATTTACCATATCATCTACAGCAACATATATGACCAACGACATATTAAG 829
 Db TTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 883
 QY 830 GATTTATGACAAATCATGATATTTATCTTAACATTTCTACCAACATTTAGCAC 889

Db 864 GTAATTTATTTAAATTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 936
 QY 890 ATATGATACAAATTAACAAACAAACAGCATATGTTAAATTAATCTCCAAATTAATTTGT 949
 Db -TTTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 995
 QY 950 AGTAAGATACGAAATTTACATCTCAATTTCCAGTAAATTAAGAGAAATTTTAA 1009
 Db TTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 1055
 QY 1010 CAATATATCCAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 1069
 Db TAAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 1115
 QY 1070 CTTCCTGTAGATATGATCTTGTGATATTAATTAAGAACTGGTGAATCATTTATGAA 1129
 Db TTGTTTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 1175
 QY 1130 ACAATTAAGTATGATGTTCTTAAATTAATTTACCATCAAGATTTATAGGTGACAGAT 1189
 Db ATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 1235
 QY 1190 GGAGCTGATTAAGAAACATTAATGATGATGATGATGATGATGATGATGATGATGAT 1249
 Db TTTTGAATTAATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 1295
 QY 1250 GGTATATATGCTGTGCTACTGTTAACTAAAGGATGATGATGATGATGATGATGAT 1309
 Db AAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 1355
 QY 1310 ATAGCAAAAATTCATCATATGATCTTCAATCATTAATTTGCTCCATCATGCTTTGAGAT 1369
 Db TAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 1415
 QY 1370 CTAATTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1429
 Db TTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 1475
 QY 1430 AAAGTGAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1489
 Db AAATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 1535
 QY 1490 ACTGTAAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1549
 Db TTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 1595
 QY 1550 TCTATTTTCAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 1609
 Db TATATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 1655
 QY 1610 TCTAGTCTATGATCTTGTATTTTCAACAGATTAATTAATTAATTAATTAATTAATTAATTA 1669
 Db TTAATTTATTAATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 1715
 QY 1670 GAATACAGCTTCCACATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1729
 Db ATATATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 1775
 QY 1716 AATATATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 1782
 Db AATATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 1835
 QY 1730 AAGTATGATATTTTCAATTAATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 1842
 Db TAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 1895
 QY 1836 AAAAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 1895
 Db AAAAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 1955
 QY 1843 ATTAAGTAAA-----TTAACTTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 1896
 Db TAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 1955
 QY 1897 TCACTATCATGATTAATTAATTAATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 1956
 Db TAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 2012

QY	1957	ATACGATATATTTAAATTAACGTAAACACGCAATTTACACGTAATATATACATGATGCGCGA	2016
Db	2013	AAATGATATTTAAATTAATAATATTTTAAATATTTTATTTTAAATAATTTATATTA	2072
QY	2017	CATATTAATTCGTATCAATCCTTGACATATTTCAACAATAAATATG-----ATTAAGT	2070
Db	2073	AAATATTTTAAATATTAATAATTAATAATTAATAATTAATAATTAATAATTAATAATTTT	2132
QY	2071	TTTTTTATGGCCATTAGTCTCTAATTAATAATGCGAATTTACCAGAAAAAGTTAAATACA	2138
Db	2133	TTATTTTATTAATAATTAATAATTAATAATTAATAATTAATAATTTATTTTAAATTTTA	2192
QY	2131	CATTCCTGCAACATATGAAAGATCAATPACAGTCAAAATTAATTAACAAATTTCTA	2190
Db	2193	AAAAATGAAAAAATAATTAATAATTAATAATTAATAATTAATAATTAATAATTTATTTT	2252
QY	2191	ATATACACCTACTCTTAAACATAGAACCAATAGATATATCATGCTAATAGATATCC	2250
Db	2253	AAATATTTAAATTAATTTAAAAAATTTTAAATTTTAAATTTTAAATTTTAAATTAAT	2312
QY	2251	CTACAGCATGTATCAATTAATTAAGAAAAATCGTTATGTATATTTATCAATTAACAAAT	2310
Db	2313	TATTAATTTTGTATTTAAATTAATAATTTTATTTAAATTTTAAATTTTAAATTTTAAATTTT	2372
QY	2311	TTCCACTTCACACAACAAACGCCAAGTCTATTTAAACAATTCAGAAATTAATATATAC	2370
Db	2373	TAAATTAATTTATTAATTTTATTTATTTTGGAAATTTAAAAAATAATTAATTTATTT	2432
QY	2371	TTCAACAACATAGTTATCAATTCAT--AAATACATCATTCCTTTTAAAAAGATG	2432
Db	2433	TTTTAAAAATTAATTTTATTTTATTTTATTTTAAATATTTTAAATTTTATTTTGTTTAAAT	2492
QY	2428	ACTTCACATTCATATCAATTAATTAATTAACATGCAATGTTATTCAGAAACATTAACAA	2487
Db	2493	AATTAATAATTTAATGAAAAATTAATAATTAATTAATTAATTAATTAATTAATTAATTA	2552
QY	2488	TCCAAGAAATTTCAATTTACGCGTGAATGCGAAGCTTCACATCCACAAATTCAGCAGT	2547
Db	2553	TTTTAAAAATTTAAATTAATTTTAAATAATTAATAATTAATAATTAATAATTAATAA	2612
QY	2548	TCGTGATACATTAAGCTCCAAAAACGAGATACAGATTTTAAATATGCTCTTTCAAATTC	2607
Db	2613	TATAAATTAATTTAAAAAATAATTAATAATTAATAATTAATAATTTTAAATAA	2672
QY	2608	TTCAAGAGCTTTACGTAATGCTACAGAAATTAACCTTAGACCTTAACATCATTAATAGG	2667
Db	2673	ATTA--ATTTTAAATTTTAAATTAATTAATTTTATTTTAAATTAATAATTAATTTGT	2730
QY	2668	CGAATTAATCTATATCTACTGCTAAATCTTTTGGTGAATCAATCT	2714
Db	2731	TAAATTAATTAATTAATAATAATAATAATAATTAATTAATTTGTTATTT	2777

AD	RESULT 17
DA	ADA31919/c
ID	ADA31919 standard; DNA, 1557 BP.
XX	
AC	ADA31919;
XX	
DT	20-NOV-2003 (first entry)
XX	
DE	DNA encoding Acinetobacter baumannii protein #3206.
XX	
XX	des: gene; Acinetobacter baumannii, bacterial disease; antibacterial,
KM	vacaine; plant biocontrol agent.
XX	
OS	Acinetobacter baumannii.
XX	
FN	US6562958-B1.
XX	
PD	13-MAY-2003.
XX	

PF	04-JUN-1999;	59US--00328352.
XX		
PR	09-JUN-1998;	98US--0086701P.
XX		
PA	(GENO-) GENOME THERAPEUTICS CORP.	
XX		
PI	Bretton G, Bush D;	
XX		
DR	WPI; 2003-576092/54.	
XX	P-PSDB; ADA36045.	
PT	New Acinetobacter baumannii proteins and nucleic acid, useful as reagents	
CC	for diagnosing a bacterial disease, as components of antibacterial	
CC	vaccines, as targets for antibacterial drugs, or as biocontrol agents for	
CC	plants.	
XX		
PS	Example; SEQ ID NO 3206; 328pp; English.	
XX		
CC	The invention relates to isolated Acinetobacter baumannii nucleic acids.	
CC	The A. baumannii nucleic acid and polypeptides are useful as reagents	
CC	for diagnosing a bacterial disease, as components of antibacterial	
CC	vaccines, as targets for antibacterial drugs, to detect the presence of	
CC	A. baumannii and other Acinetobacter species in a sample, in screening	
CC	compounds for the ability to interfere with the A. baumannii life cycle	
CC	or to inhibit A. baumannii infection, and as biocontrol agents for	
CC	plants. The present sequence represents DNA encoding an A. baumannii	
CC	protein.	
SQ	Sequence 1557 BP; 481 A; 321 C; 330 G; 425 T; 0 U; 0 Other;	
XX		
Query Match	3.4%; Score 160; DB 8; Length 1557;	
Best Local Similarity	48.3%; Pred. No. 2,4e-16;	
Matches 553; Conservative	0; Mismatches 575; Indels 18; Gaps 3;	
OY	2639 TTTCGCTGAATCACTCCTTAAGGATTGTACTAATTTAAATCCATTCTATACTTACCTACGT	2755E
Db	1208 TCCTTGCGATTAACCTCACGCCAGCATATCATGTAGCGGTAAACAAGCGACGGATTTTCA	1149H
OY	2759 AATTGCTTAAGTACCCTGCGCATTTCATATGATGATGCTATTTAAAGCTTCATCACTT	2818L
Db	1148 AGCTGACGACATATACCGAGCGCATGTAAATGTTGCATATAGTCCGTTGTTACGTGCGT	1089S
OY	2819 AATACATGAGTACACAAACCATTTTTTTTCAAGATGCTATTCAAATAATCCNACACAT	2878R
Db	1088 AAAATCTTTGACTTAGCGCCCTAGCTCTTTACCCGCGGAGCCAAAGAATGTTGACCAAC	1029D
OY	2879 TCCGGTATATCTGTACAAATATTTACTTAAAGACGGTACTGTAAATGGAGACATTTAAT	2938L
Db	1028 ATCCGGATATGCTTCACTGCGCATGCGCAAGCTTGGAAATATGAATTCGAATTAACATTAGA	969
OY	2939 CTAAATATTAATCCGACAAAACCTCCAGAGCTTTTACTTCACCTTCAATATCTTTGGA	2998B
Db	968 CGGTGTATATAGTITTTACGGAACGACCTTCACTTCCACGTTTTCATAGCTTTGATGG	909
OY	2999 GAAGACACAATTAATTCACATCTTACTTAACAAGAACTTACTATTTTCCCTATATAT	3058R
Db	908 GTTCCCGCGCAATTTCTGACATCTTAACCGGTATATGTCCACCAACTCGGTAAAC	849
OY	3059 TTTCGCTTCGTATATATCTGAGTAATCTTAATTTGGCTATCGTATATCTGTACT	3118L
Db	848 TCACCATCTGCGAGACACAGTACACATGAGATTGTGTTTCAAAAGGSCATGTCCCAATT	789
OY	3119 TCATCTTAATAATAACGTACACATGTTGCTTGTCTTAATAATTCGAATATGAGAGGACT	3178H
Db	788 TCATCAAGAAATAGTGTTCGCGCATTTGCTTTCAAGCGGCTTGATGTGGGTGTTT	729
OY	3179 CTATGACACAATATATTAATTAATCTTCTCACTAACCAATATATTAACCAAGTATATTTA	3238R
Db	728 GCACCCCGTAA-----AGCACCTTTTCATGACCAAAATATTCGTTTGATTAATCT	675
OY	3239 GCTGTGTAGCATAGTAGTACATATGATATTAATGAGATATCAATCCCCTTGATTTTTTA	3298R
Db	674 TTGGAATTTGCTGCGCATGTTTAAAGCATATAAGGCTTAGCGGCTCTGTGTATATGTTTA	615

QY 924 GTAAAAATCTGCAATATTTATTTGTAAGTAAGTATCGAATTTACATCAGTCAAAATTC 983
 DB 44213 AAAAAATCTACTCTCAAAAAATCAATCCACTGATAAAAAATTAACATTAACCTAA 44154
 QY 984 CCAAGTAAATAGCAGAGAAATTTTCAATATATCAATTTTATCTCTGCTGCA 1043
 DB 44153 ATATCTCCCATACATATATTTTCACTACATCAAAAATTTTAAACAAAAATTA 44094
 QY 1044 GTTTTGCTAAAGAAATTTGAGAACATCTTCTGT--AGTATAGACTTGCTGGTATTA 1101
 DB 44093 ACTAAAAATTAATTAATTAACCACTTATTTCAAAAAAATTAACCAACAAAAA 44034
 QY 1102 TAAAGAACTGTGTAATCATGTAGAAAAGTAAAGTAATGATGTTCTAAAAATTAATA 1161
 DB 44033 AAAAAAATTAATTAACCAATATTAATTAATTAACAAAAAATTAATTAACATA 43974
 QY 1162 CCATCAAGATATTAAGTGTACAGATTGAGCTGCATTTAAAGACATATTTGCAATTGC 1221
 DB 43973 CATTAATAATTTTAAATAATTTACA--AAAAATCTCTCTTAACCAAACTTTACACATTAA 43915
 QY 1222 ATGTGAATTAATCGCTGAAAAAATTTAGTAAATATGCTGTGCTACTGTTATTAACATA 1281
 DB 43914 AAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 43855
 QY 1282 AGGCAATGAATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1341
 DB 43854 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 43795
 QY 1342 ATTAATGTCATCATGCTGTCGAGATCTAATTAATTAATTAATTAATTAATTAATTAAT 1401
 DB 43794 AACAAATTAATCTTTATTTATTTATTTATTAATTAATTAATTAATTAATTAATTAATTA 43735
 QY 1402 CAATATGCTTTGAGCTGAATTAAGAAAGTGAATTAATTAATTAATTAATTAATTAATTA 1461
 DB 43734 ATTAATTAATTTTAACTTTAA--AACTATTAATTAATTAATTAATTAATTAATTAATTA 43679
 QY 1462 CAACCTAAAGCTTTGTAAGAACAGTACTTAATAACCATGATATCATTAGCAAAAAA 1521
 DB 43678 CTATCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 43619
 QY 1522 ACTTAATGTAAGTCAATTTGCAATCTATTTTCAATTTATTAATTAATTAATTAATTAAT 1581
 DB 43618 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 43559
 QY 1582 ACTAGTAAGCATATCAACATTTTACTAGCTATCACTCTTGAATTAATTAATTAATTAAT 1641
 DB 43558 AAAATGGAATTCACATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 43499
 QY 1642 CAATGTAATTAAGGCTCAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1701
 DB 43498 CCAACTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 43439
 QY 1702 GTCAATATCCATGCTTAATCACTCACTAAGTATTAATTAATTAATTAATTAATTAATTAAT 1761
 DB 43438 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 43379
 QY 1762 TATTAATTAATTTCTAGCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1820
 DB 43378 CATTAACGAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 43319
 QY 1821 ACCAATGTTTATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1880
 DB 43318 ATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 43259
 QY 1881 ATGTTTGCACACTTTTCTACTACATCATGTATTAATTAATTAATTAATTAATTAATTAATTAAT 1940
 DB 43258 CAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 43199
 QY 1941 CTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2000
 DB 43198 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 43139

QY 2001 TACATAGATGCTCGCATATTAATTAATTTCTATCATCTTGCATATTTCAACATAAATA 2060
 DB 43138 TTACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 43079
 QY 2061 TGAATTAATGTTTTTTTATGCCCCATTAATTTCTTAATTAATTAATTTGCTAATTAACCAAAAG 2120
 DB 43078 AAATTAACAAAAAATCTTCTCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 43019
 QY 2121 TTAATAATCACTCTCGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2180
 DB 43018 TCCATAAATAATTAACAAAAACCAACCCCAACAAACAAACCAACCAACCAACCAACCAACCAACCAAC 42959
 QY 2181 CAATTTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2240
 DB 42958 AACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 42899
 QY 2241 TAAAG--TATCCCTACAGCATGTATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2299
 DB 42898 TTATTAACAAATCTTAACAAATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 42839
 QY 2300 AATTAACAAATTTTCCAACTTCAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 2359
 DB 42838 CATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 42779
 QY 2360 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2419
 DB 42778 ATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 42719
 QY 2420 AAAAGATGAATTCACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2455
 DB 42718 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 42683

RESULT 19
 AAS46530/c
 ID AAS46530 standard; DNA; 7814 BP.
 XX
 AC AAS46530;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Tumour suppressor gene derived chemically modified sequence #252.
 XX
 KW Human; tumour suppressor gene; oncogene; antitumour; cytostatic; cancer;
 KW tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
 KW cytosine methylation; ds.
 XX
 OS Homo sapiens.
 XX
 PN W0200168912-A2.
 XX
 PD 20-SEP-2001.
 XX
 PE 15-MAR-2001; 2001MO-EP002955.
 XX
 PR 15-MAR-2000; 2000DE-01013847.
 PR 06-APR-2000; 2000DE-01019058.
 PR 07-APR-2000; 2000DE-01019173.
 PR 30-JUN-2000; 2000DE-0103529.
 PR 01-SEP-2000; 2000DE-01043826.
 XX
 PA (EPIC-) EPIGENOMICS AG.
 XX
 PI Olek A, Pieperbrock C, Berlin K;
 XX
 DR WPI; 2001-602752/68.
 XX
 PT Fragments of chemically modified genes associated with tumor suppressor
 PT genes and oncogenes, useful in designing primers and probes for analyzing
 PT diseases associated with cytosine methylation state e.g. cancer.
 XX
 PS Claim 1; SEQ ID NO 252; 27bp; English.

Query Match	3.3%;	Score 155.2;	DB 4;	Length 7814;
Best Local Similarity	44.0%;	Pred. No. 1.4e-15;		
Matches 945;	Conservative	0;	Mismatches 1188;	Indels 17; Gaps 6

[illegible][illegible]

QY 1302 CACTATATATAGCAAAAAATCATCATAGATCTTCATCATTAATGGTCATCATGTC 1361
 Db 728 GCTTGTGTGCGGATGGGT-----GCCATCTCTGACATTTATCGTTGAAGTG 781
 QY 1362 TTGGAGATCTAATATATACATGACAGAACATTCGCGCAATAGCTTTGGACTAG 1421
 Db 782 TGGGAGATCTGATGCTTACTTGTACAGTGTTCATTTCTGCAATGGCGGTAACT 841
 QY 1422 AATAGCAAAAGGTAGCAAAATATTAATATCATTAATGATCAACCACTTAAGCTTGAAG 1481
 Db 842 TGTAGAGACAGGACATAGAGTGAAGAGTCTGAAAACATGGAAATGGCTTGAAG 901
 QY 1482 GAACAGTCTGTGAAAACCATGATATCTTAGCAAAAAAATTATATGAGAACCA 1541
 Db 902 GTGTGCAACGACAAAAGACAGATGAGCTGACCTCAACGCTAGGTGTGAGATGCCA 961
 QY 1542 TTGCATATCTATTTTCAATTTTATCATGAGATATATCACTAGATTAAGCCATATCA 1601
 Db 962 TTACTCAAAATCTATATGTTTGTATATGAGAAAGATATCAAAAAAGACGAAAG 1021
 QY 1602 ACATATAT 1610
 Db 1022 AATCATGT 1030
 RESULT 22
 ABL33396/C
 ID ABL33396 standard; DNA; 5641 BP.
 XX ABL33396;
 AC
 DT 26-MAR-2002 (first entry)
 XX
 DE Human immune system associated gene SEQ ID NO: 1369.
 XX
 KW Human; immune system disease; cytosine methylation; antiscismatic;
 KW antileukemic; anti-HIV; anticonvulsant; ophthalmological;
 KW antineoplastic; antidiabetic; antidiabetic; antipsychotic;
 KW antineoplastic; cancer; eye disease; arteriosclerosis; anaemia;
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
 KW ds.
 OS Homo sapiens.
 XX
 PN WO200200928-A2.
 XX
 PD 03-JAN-2002.
 XX
 PE 02-JUL-2001; 2001WO-EP007537.
 XX
 PR 30-JUN-2000; 2000DE-01032529.
 XX
 PR 01-SEP-2000; 2000DE-01043826.
 XX
 PA (EPig-) EPIGENOMICS AG.
 XX
 PI Olek A, Pieperbrock C, Berlin K;
 XX
 DR WPI; 2002-130909/17.
 XX
 PT Nucleic acid comprising fragment of chemically modified gene, useful for
 PT diagnosis and treatment of diseases associated with abnormal cytosine
 PT methylation.
 XX
 PS Claim 1; SEQ ID NO 1369, 32pp + Sequence Listing; German.
 CC The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid

CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 CC diseases. The present sequence is a gene of the invention
 XX
 SO Sequence 5641 BP; 1442 A; 34 C; 930 G; 3235 T; 0 U; 0 Other;
 Query Match 3.3%; Score 155; DB 6; Length 5641;
 Best Local Similarity 43.9%; Pred. No. 1.5e-15;
 Matches 1027; Conservative 0; Mismatches 1280; Indels 34; Gaps 7;
 QY 176 ATTTTCATATCTCCAAATTTTAACTGACGCAAGATTTATTAATGACAGAACCCCTG 235
 Db 4946 ATTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4887
 QY 236 TACTCTTATGCAAGAACTCTATTAACCTTATCAAGGCTCTTACATAGAACAACACTGTAT 295
 Db 4886 CACCCAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 4830
 QY 296 TCAAGAAATACCTCTCTCTGTTTCATCCATATTTTCCAAATTTTACAAATATCATAGA 355
 Db 4829 ACGATATATATTTTCACTACTTCAAAATTAATTAATTAATTAATTAATTAATTAATTA 4770
 QY 356 GTTCTTATTTGACATACATATTTAAATCAATCAAAACAAATATATATATATATAT 415
 Db 4769 TAAATCTTATTTTCAATCTATCATCTTACATATTTTAAATATATATATATATATAT 4710
 QY 416 AATTAATTAATTTTCAATTTTGT-----CTTAGAGAGACATCAATTAATGAGACCTA 471
 Db 4709 ATTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4650
 QY 472 ACTTATTAATTAATTTCAATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTA 531
 Db 4649 AAAAAAAAAAAAAAAAAATTAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTA 4590
 QY 532 TGACATGATTAACATATAGTACACATTTCTTAACTAGATTAATTAATTAATTAATTAAT 591
 Db 4589 ATAACTCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4530
 QY 592 ATATTTTGTGCTGTTCTTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 651
 Db 4529 ATATTTTAAATTAATTTTCTTCTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 4470
 QY 652 TGCAGATCTATTTGGACAGCAATAGCAATGACATGACATGACATGACATGACATGACAT 711
 Db 4469 TAAACCTCAATTAATTAATTTGCAAAATTTCTTAAATTAATTTCTTCAATGACATTA 4410
 QY 712 CTATGGGAGCGATCATATGAATATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 771
 Db 4409 ATTAAT-----AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4357
 QY 772 ATATTTTACCAATATCATCTTACAGACATATATGCAACGACATATATGACAGAAAT 831
 Db 4356 AACAATTAACACTCAAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4297
 QY 832 ATATCTGACAAATATATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 891
 Db 4296 ATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4237
 QY 892 ATGTACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 951
 Db 4236 ATTAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4177
 QY 952 TAAAGGATGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1011
 Db 4176 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4117
 QY 1012 ATATATCAATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1071
 Db 4116 CTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4057
 QY 1072 TCTTGTAGATATGATCTGCTGATATTAATTAATTAATTAATTAATTAATTAATTAATTA 1131
 Db 4056 AATCTATCTTAATTTCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3997

QY	1132	AATAGATGATGATGCTTCTTAAAAATATATATACATCAAGATATATAGGTGACATGCG	1139
Db	3996	AAAAATATACAAAATATATCTTTCACATCAAAATATTTAAAAATCTTAAAAATATATATA	3993
QY	1192	AGCTGCATTAAGAACATATATTCGCAATTCGATGTGGAAATATCGGTGAAAAAATTAGG	1251
Db	3936	TAAATTTAAAAATAAAACAATCAATTAATAAACTTCACAAAATTAACAAAAACGTAAATTT	3877
QY	1252	TAAATATGCTGTGTCTACTGTTATATCTAAAGGCAATGATGAATTAATAACCTATAT	1317
Db	3876	AACAAATTTATTTAATATACATATATATTAACCGATTAATTTTATATCTTATCAACACATTTA	3811
QY	1312	AGCAAAAAATCATTCATATGATCTTTCATPACATTAATGGTCCATCATGTCTTGAGATCT	1371
Db	3816	ATTACCAATTAAMAACTCAATATTCMAAAAAATTAACATTA-----AATAACT	3767
QY	1372	AATATTAACATGACACAGAACTTTCACGCAATATGGCTTTGGACTAGAAATGAGAA	1433
Db	3766	ATATATTTAATTTGCAAAAACTTAATTCGACAAATTTTTATATCTTATCAACACATTTA	3707
QY	1432	AGGTGAAAATATATACTATATATGATCAACAACCTAAAGCTGTGAGGAACCAATGAC	1491
Db	3706	AAACATATTAATAAATAAATACTATTAATATACCTT---TTTTCTTCGAAACCTAA	3651
QY	1492	TGTAAACCACTGAATCATATGACAAAAAACTTAATGTAGAACACCAATTTGCATTC	1551
Db	3650	TATATATTAATAATTAATATTTTAAAAATCMAAAAAAATAAATATATATATATAA	3591
QY	1552	TATTTACAATTTATTAATCATGGAATATATCACTAGATPAAGCCATATCAACATATATC	1611
Db	3590	AATATAAAAAATAATAATATATATTAACCACTAATATGATCAAAAAAATAATTAATC	3531
QY	1612	TTAGTCATCATCCCTTGATATATTTCAACACATGATTAATPAAGCCMAAAATATATGA	1671
Db	3530	TCTTCGATACTACTATCAACATTCATCAATATATTAATAAATCAATTAATTAATATCC	3471
QY	1672	ATACAAAGCTTCAACCATATATTAAGAAGCTCA---ATPCCATCGTTACATCCAA	1727
Db	3470	AATPACCAACCCCATTAATAAATAAACTAATACCCTTAATATTTACATTAATATCTAA	3411
QY	1728	CTAAGTATGATACATATATACATTAACAATTAAGTATATCTTAATTTCTAGTATAT	1787
Db	3410	TATCCATATTCGCAACAAATTTTATTAACATTAATAAATAAATAAATATATATCCATATA	3351
QY	1788	ATATPAACAAATTCCTATCTCCATCTCAATTAACCAATGTTATCAGATTAATAATAG	1847
Db	3350	AAAAAAAATCAAAACMAAAACCTACCTATTAACACATATTAATCAACCTCTTTAA	3291
QY	1848	TCAATATTAACCTTTACTATTTTATTAATATATATGTTTGACACCTTTTCACTACATCA	1907
Db	3290	AAAAACCTCAAAATPACATTAATAATTTTTCACAAACCTAAAAAACAATAATTAA	3231
QY	1908	TGATATATATGAACATAATTTTTCATAAAATPACTATTAATAAAGAACCAATPACATATTT	1967
Db	3230	AAAAATAAAAAAAATACATATAAAAATATCAATACAAATAAAAAATATCAATPAAAAAAAT	3171
QY	1968	AAATTTACTAAACACCTATTTACTACGTAATATTACA--TGATAGTCTGACATATATAA	2025
Db	3170	AAATTAATAAAAAACCAATTAATAAATTCCTACATCAAAATATACAAATCTAAATATAAA	3111
QY	2026	TTCTATCATCTTGACATATTTCAACATTAATAATATGATTAACGTGTTTTTATAGCCAT	2085
Db	3110	TTTATATATATAAATTAATATATATATTTAACTACAAAAAAAATAACATCTTTTAA	3051
QY	2086	TAGTCTTAAATTAATGCTAATTAACCCAAAAAGGTTAAATATACATCTCTGCACATAT	2145
Db	3050	TAAATTTATATCAAACTAAAAAATTAATAAAAAAATAAATAAATAAATAAATAA	2991
QY	2146	ATGAAGAAGATACAAATACAGCTCAAAATTAATAACAAATTTCTATATTAACGACTACT	2205
Db	2990	CTCAAAAAATATATAACCACTTAAACATACCAACACATPAATAATATCAAAATCTAA	2931

Accession	Source	Strain	Sequence	Position
AY	CTAATACCTTACGACCAATAGTATATCATACCTTAAGTAGATATCCCTACGACATGATCA	2265		
Db	2230 AACGAAAAAAAAAAAAATCAAAAAAAAAAAAAATCCAAAAAAAAAAATACCTAATAA	2871		
Oy	2266 AATATATTAAGAAAAATCGTTATGTATATTATATACATTAACAAATTTCCAACTTCAACAA	2325		
Db	2870 ATTATATTTAAAAAACTACACATATTAATAAAAACTGCACAACTTTTAAACAAATTAATCAAA	2811		
Oy	2326 CAATACCTCAAAAGTCTATTTATTAACAAATTCAGATATATATATCTACCAACACATAG	2385		
Db	2810 AAAAATTCACAAATTAACACATCATATATTAATAAAAACTTAATAAAAAACAAAAA	2751		
Oy	2386 TTATCAATTCCTATATATACATCATATCTTTTAAAAAGATGACTTCACATTCATACAT	2445		
Db	2750 TCTTAAAAATCAACAAAAATTAATAAAATTTTACTTACAAAAAAACCTTAATATATTAAC	2691		
Oy	2446 AATATATTTATACATGCATGTTTATTCAGAAACATTCACATTCACAAAGAAATTTCAATT	2505		
Db	2690 AACCAATTTCTCAACAAAAACAAAAAAACCTCCAAAAACATTAATATACATATTCAAAT	2631		
Oy	2506 A 2506			
Db	2630 A 2630			

RESULT 23
AAV74824/C
ID AAV74824 standard; DNA; 918 BP.

AAV74824;
16-MAR-1999 (first entry)

Staphylococcus aureus contig SEQ ID #513.

Computer readable medium; vaccine; S.aureus infection; immunodetection; cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy; skin infection; surgical wound infection; scalded skin syndrome; toxic shock syndrome; ds.

Staphylococcus aureus.

Key	Location/Qualifiers
FT misc_feature	781..840
FT	/*tag= a
FT	/note= "these bases represent a line of missing text in the sequence listing in the specification. They are included to maintain the nucleotide numbering given in the specification for this DNA sequence"

EP786519-A2.

30-JUL-1997.

07-JAN-1997; 97BP-00100117.

05-JAN-1996; 96US-0009861P.

(HUMA-) HUMAN GENOME SCI INC.

Kunsch CA, Choi GH, Barash SC, Dillon PJ, Fannon MR, Rosen CA; WPI; 1997-374922/35.

Polynucleotide(s) and proteins derived from Staphylococcus aureus - stored on computer readable medium and used in the production of anti-S. aureus vaccines.

Claim 1; Page 1440-1441; 3271pp; English.

CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
 CC the S.aureus DNA sequences allows putative functions to be assigned so
 CC that protein-encoding or regulatory regions of commercial, therapeutic or
 CC industrial importance can be obtained. Specifically, sequences which are
 CC likely to encode antigens have been identified and these polypeptides can
 CC be used in a vaccine composition against S.aureus infection. The
 CC polypeptides can also be used in a kit for the immunodetection of
 CC S.aureus in a sample. S.aureus is implicated in numerous human diseases,
 CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,
 CC skin and surgical wound infections, scaled skin syndrome, toxic shock
 CC syndrome, etc. Organisms transformed with the DNA sequences can be used
 CC for recombinant production of the polypeptides. The new DNA sequences
 CC (and their fragments) are useful as primers or probes for isolating
 CC homologues of any of the S.aureus DNA sequences contained on the computer
 CC readable medium

XX Sequence 918 BP, 271 A, 177 C, 113 G, 297 T, 0 U, 60 Other;

Query Match 3.2%; Score 154.2; DB 2; Length 918;

Best Local Similarity 53.6%; Pred. No. 1.9e-15;

Matches 346; Conservative 0; Mismatches 293; Indels 6; Gaps 1;

QY 1022 ATTTTATCTCTCTGTCAGATTGCTTAAGAAATTCGAAATCTTCTTGTACT 1081
 DB 732 ATTGGCGTGTGTGTCAGGCGCAATCATGGGAAGAGTGTGTCAGCAACCACTCA 673
 QY 1082 ATAGTACTGCTGTGATATTAAGAACTGGTGAATCATGATGAAACAATAAGTAT 1141
 DB 672 GTTGCGCTTCATCAAAAGATTAAGTGAATTAATTAACCAAGATTTATTTATGAT 613
 QY 1142 GATGTTCTAAATAATATATACATCAAGATTTATAGTGTACAGATTGAGCTGATTA 1201
 DB 612 GATATTTCCGTGTGACAGATGATGATGATGATGATGATGATGATGATGATGAT 553
 QY 1202 AAGACATATATGCAATTCATGCAATTAATGCGTGAAGAAATTTAGTAAATAGCT 1261
 DB 552 AAAAAATCATCGCAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAG 493
 QY 1262 GTTGCTACTGTTATTAATTAAGCAATGAAATTAATAAATTAATTAATTAATTAAT 1321
 DB 492 AAGCGTCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 439
 QY 1322 CATTAATGATCTTATCATTAATTTGTCATCATGCTGAGATCTTAATTAATTAACA 1381
 DB 438 TTATGTCGCGATCTATGATCATTTCTAGTTTATGATGATGATGATGATGATGAT 379
 QY 1382 TGTACACAGAAATTCACGCAATATGCTTTTGAATGAAATGAGAAAGTAAAT 1441
 DB 378 TGCACATCAACATTCGGAATTCACATTAAGATTAATTAATTAATTAATTAATTA 319
 QY 1442 ATTAATACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1501
 DB 318 ATGACATCAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 259
 QY 1502 CTGATATCATTAAGCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1561
 DB 258 GTTATCATTTAGTAAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 199
 QY 1562 TTATTAATGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1621
 DB 198 GTATTAATTTGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 139
 QY 1622 ATTCCTTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1666
 DB 138 AATTCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 94

RESULT 24
 ABL34155/c
 ID ABL34155 standard; DNA, 15548 BP.
 XX
 XX ABL34155;
 XX AC
 XX

DT 26-MAR-2002 (first entry)

XX Human immune system associated gene SEQ ID NO: 2128.

DE Human; immune system disease; cytosine methylation; antiasthmatic;
 XX antiarteriosclerotic; antiasthmatic; cytosine; noctropic;
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
 KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
 KW antileukemic; cancer; eye disease; arteriosclerosis; anaemia;
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
 KW ds.

OS Homo sapiens.

XX WO200200928-A2.

XX 03-JAN-2002.

XX 02-JUL-2001; 2001WO-BP007537.

XX 30-JUN-2000; 2000DE-01032529.

XX 01-SEP-2000; 2000DE-01043826.

XX (EPIG-) EPIGENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K;

XX WPI; 2002-130909/17.

XX Nucleic acid comprising fragment of chemically modified gene, useful for
 PT diagnosis and treatment of diseases associated with abnormal cytosine
 PT methylation.

XX Claim 1; SEQ ID NO 2128; 32pp + Sequence Listing; German.

CC The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 CC diseases. The present sequence is a gene of the invention

XX Sequence 15548 BP; 4209 A; 247 C; 2903 G; 8189 T; 0 U; 0 Other;

Query Match 3.2%; Score 153; DB 6; Length 15548;

Best Local Similarity 44.7%; Pred. No. 3.3e-15;

Matches 981; Conservative 0; Mismatches 1185; Indels 27; Gaps 9;

QY 237 ACTCTTATGACAGAACTCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 296
 DB 13534 ACACATACGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 13475
 QY 297 CAAGATATACCTCTGCTGTCATTCATTAATTTCCATTAATTTATTAATTAATTAATTAAT 356
 DB 13474 CATTAATACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 13416
 QY 357 TCTCTTATGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 416
 DB 13415 ATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 13356
 QY 417 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 476
 DB 13355 ATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 13296
 QY 477 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 536
 DB 13295 CTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 13236
 QY 537 TGATTAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 596

Db	1847	AAAAAAAAAAAAAAAAATTAATTAAATTATTTTATTTTATTAATAAATAAATAA	1906
Qy	1544	TGCATATCGATTTACAAATTATTAATCAATGAGAAATA-TCACTAGTAAAGCATATCCAA	1602
Db	1907	ATTTTATTAATAAAAAAAAAATTAATTAATTAATAATAAATAAATAAATATTTTAA	1966
Qy	1603	CATATTAATCTTAGCTATCATCCCTTGATTAATTCAACACATGATAATTAACGCTCAA	1662
Db	1967	AATAAATAATTAATTAATTAATTTTATTAATAAATTTAAATAATTTAAATTAATAA	2026
Qy	1663	AAATTAAGATCAACGCTCACACATATTAATAAAGACATACCCTCGTTACA	1722
Db	2027	TAAATTAATTAATTTAAATATTTTATTTTAAATAATTTTAATAAATATTTTAAATA	2086
Qy	1723	TCCAACTAGTAGTACATATTAATCAACAACCTAAGTATTAACCTAATAATTCCTAGC	1782
Db	2087	TTACAAATTAATA-ATTATTAATAATTTATTAATTAATTTAAATAATTTTATTTAATAA	2145
Qy	1783	TAAATTAATAACAATTCCTATCCCTATCTCAATTAACAATGGTATACGATTATA	1842
Db	2146	AATTAATAATTAATAAAAAAAAAATTAATAATTTATTAATAAATTTAAAAAACAATAAA	2205
Qy	1843	ATAAGCTAAATTAACCTTTTACTATTAATTAATAAATAATATAGTTGCACTTTTCACTA	1902
Db	2206	AAAAATTAATAAAAAAAAAATTAATTTTAAATAAATAATTTTAAATTTTAAATAATA	2265
Qy	1903	CATCATGTATTAATGACATTAATTTTACATTAATACTAATAAAGAACCATACAT	1962
Db	2266	AATTAATAAATAATTTAAATTTTATTAATTAATTAATTTTAAATAATAATTTATTAATTTT	2325
Qy	1963	ATATTAATTAATTAACCACTATTTTACTAGTATTAATTAATTAATAGATGCTCCGACATAT	2022
Db	2326	ATTTAAATTAATAATTTATTAATTTTAAATAATTTTAAATTTTAAATAATTAATTTAT	2385
Qy	2023	AAATTCATCATCCCTGACATTTTCAACAATAAATATGATAACGT-----TTTTT	2076
Db	2386	AAATTTATTAATTTTTCGAAAAATTAATAAATAATTAATTTTATTTTAAAAAATTAAT	2445
Qy	2077	TATGCCATTAAGTCTTAATAATAAATGCTAATTAACCAAAAGGTTAAATATACATTC	2136
Db	2446	TTTTTTATTTTATTAATTAATAATTTTATTAATTTGTTTAAATAATTAATAATAATA	2505
Qy	2137	TGCAACAATATGAGAGAGATCAATACGCGCAATAATAATTAACCAATTTCTAATATA	2186
Db	2506	ACGAAAAATTAATAAATAATTAATTAATAATAATAATAATAATAATAATAATAATAATA	2565
Qy	2197	CCACTTACTTAACATAGAACCAATAGTATCATATCACTAAGTATAGATATCCCTACAG	2256
Db	2566	ATTAATAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAATA	2625
Qy	2257	CATGATCAATTAATAAGAAAAATCGTATGCTATTTATTAACATTAACAATAATTTCCA	2316
Db	2626	AAAAAAAAAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA	2685
Qy	2317	CTTCACAACAAAACCTCAAGTCTCATATTAACAATAATCAAGATTAATATTAATCTCAC	2376
Db	2686	TTTTTTTAAATTTATTAATTTTATTTTAAATAAATAATTCGTTAAATTAATAATA	2745
Qy	2377	AACCATAGTATCAAAATCTAATAATACATATCCCTTTAAATAAGATGACTTCACA	2436
Db	2746	TAAATAATTAATAAATAATTAATGCTTATTTTAAATAAATAATAACAATTCGTTAT	2805
Qy	2437	TTCAATACATAATAATTAATACATCAATGTTTATTCGAAAAATTAACAATTCAGAAAT	2496
Db	2806	TAAATAGTATTAATTAATAAACAATATTTTATTAATAAATAAATAATTCGAAATTAAT	2865
Qy	2497	TTT 2499	
Db	2866	TTT 2866	

ABL54314/C	ID	ABL54314 standard; DNA; 6314 BP.
XX	AC	ABL54314;
XX	DT	29-JUL-2002 (first entry)
XX	DE	Chemically treated apoptosis gene complementary to gene #7.
XX	KM	Apoptosis; HIV; Bloom syndrome; cardiopathy; neurodegenerative disorder; Hepes simplex virus; renal ischaemia; amyotrophic lateral sclerosis; cancer; ds.
XX	OS	Unidentified.
XX	PN	WO200177164-A2.
XX	PD	18-OCT-2001.
XX	PF	06-APR-2001; 2001WO-EP003969.
XX	PR	06-APR-2000; 2000DE-01019058. 07-APR-2000; 2000DE-01019173. 30-JUN-2000; 2000DE-01032529. 01-SEP-2000; 2000DE-01043826.
XX	PA	(EPIG-) EPIGENOMICS AG.
XX	PL	Olek A, Piepenbrock C, Berlin K;
XX	DR	WPI; 2002-017444/02.
XX	PT	Chemically modified sequences of genes associated with apoptosis are useful to determine methylation patterns of genomic DNA samples for diagnosis of associated diseases such as cancer.
XX	PS	Claim 1; Seq ID #14; 24bp; English.
XX	CC	This invention relates to chemically pre-treated DNA of genes associated with apoptosis. The nucleic acids are used to allocate patients for specific therapy for HIV infection, Bloom syndrome, cardiopathy, aging, neurodegenerative disorders, Hepes simplex virus infection, renal ischaemia, amyotrophic lateral sclerosis, solid tumours and cancers. This nucleic acid sequence represents a chemically treated apoptosis gene. Even SEQ ID numbers are the complementary DNA strands to the odd SEQ ID numbers. The sequence data for this patent is not represented in the printed specification but is based on information supplied by the European patent office
XX	SO	Sequence 6314 BP; 1514 A; 203 C; 1334 G; 3263 T; 0 U; 0 Other;
XX	Query Match	3.1%; Score 148.8; DB 6; Length 6314;
XX	Best Local Similarity	44.9%; Pred. No. 1.5e-14;
XX	Matches 1058; Conservative	0; Mismatches 1267; Indels 29; Gaps 12
QY	170	CAAACCAATTTTCCTCAATCTCCCATTTTAAATGAGATTATTTAATGTAAACAGAAC 229
DB	5335	CAACCATTTAAATAATTAATAAATCTAATATTAATTAATTAATTAATTAATTAATTAATTA 5276
QY	230	ACCCTGACCTCTTAATGACAGACTTAATACTTAATCAAGCTCTTAAACAATGAAGAACA 289
DB	5275	AACCTAATACATTTCTAATAATCTAAACATACATTAATAATAATACTTAATAATACTTA 5216
QY	290	CTGATATCAAGATACCTCTCTCTTGTCATCCATATTTTCCAATTATTTTCAATAT 349
DB	5215	CCTTAT----AAATTATATTCATATTAATAAAAAAACAATAAATAAATAAATAA 5160
QY	350	CATAGACTTCTTAATGACATACATTTAAATCAATCAACAAAATAATATAGACTAT 409
DB	5159	TATATTAATAACATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 51000
QY	410	CTAATTAATTAATAATAATCTCAATAGCTTTAGCAGAACTCAATTAATTAATGAACAT 469

D 5099 TTAGCAAAAAAATTAACAAAAAACCCTCTTAATAAATTAACCTTAACACT 5040
Q 470 TAACTATTAATATATTTCAAAATTAATTAAGTACTCATCATTTAGCATATATCA 529
D 5039 AAA--AAATTAATAAATTAACCAATTAACCAACCAACCAAAATTTCTTCATTCATA 4983
Q 530 TCTGACATGATTAACATAGTAACATTTCTTAACATAGATATATATCAATATAT 589
D 4982 CCTAAATCTTTCTTAATACAAACAAAAAACAATTAACCTTAACCTTAATAA 4923
Q 590 ATATATATTTTGAATGCTTTCTTAATATATTAATTAATTAATTAATTAATTA 649
D 4922 TTTAATCTCTCTCTTACGCTATTAAT--AAATACCAAAAAAACAATTTAATATAT 4866
Q 650 GGTGACGATCATTTGGACAGCAATAGCAATGCACTGTCGACATGATATATCAGTT 709
D 4865 AAAACCTACAAAAAATTAATATTTCTTAATAAATTAATTAATTAATTAATTA 4806
Q 710 AACTATATGGGACGATGATATGAATATTAACATATTAACAATTAACCTTAATAA 769
D 4805 TAAATATATTTTAAACACAAAAAATATATATATATATATATATATATATATAT 4746
Q 770 AATATATTAACCAATATATCTTACGACAAATATATGACACGCAATATGACGAA 829
D 4745 TATTAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 4686
Q 830 GTATATATGACAAAT 889
D 4685 ATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 4626
Q 890 ATATGATACAAATATCAACAACAACGATATGTTAAATAATCTCAATATATATTTGT 949
D 4625 TTAATCTCTATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 4566
Q 950 AGTAAAGATGGAATTAATTAATCACTCAATTTCCGATGAATTAAGCAAAATTTTA 1009
D 4565 ATTAAT 4506
Q 1010 CAATATATCAATTTTATATCTCTGTCGCAATTTGCTTAAGAAATTCAGAAAT 1069
D 4505 TAAACATCAAAATCTATTTACCAAAATTAATAAATAAATAAATAAATAAATAAATA 4446
Q 1070 CTCTGATGATATGATCTGCTGCTGATATATTAAGAACTGCTGATATATATATGA 1129
D 4445 AATTTATATCTTTAT 4386
Q 1130 ACAAT 1188
D 4385 ATTAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 4326
Q 1189 TGGAGCTGATTAAGAAATATATGCAATGATGATGATATATCGTGAATAAATTT 1248
D 4325 AATTAATCTCAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 4266
Q 1249 AGTATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1308
D 4265 AAAAATTAACCAAT 4206
Q 1309 TATACCAAAATATCTTCAAT 1368
D 4205 AAAAATTAACCAAT 4146
Q 1369 TCTAT 1426
D 4145 TCTAT 4086
Q 1427 GGAAT 1486
D 4085 CGAAT 4026
Q 1487 AGTAT 1543
D 4025 AAT 3966

Q 1544 TGCATATCTATTTAT 1603
D 3965 TAAATAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 3906
Q 1604 ATATATATCTTAT 1663
D 3905 ATTAATTAACCAAT 3846
Q 1664 AT 1723
D 3845 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 3788
Q 1724 CCAAT 1783
D 3787 CAACCAACCAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 3728
Q 1784 AATTAATTAACCAAT 1843
D 3727 AATCAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 3673
Q 1844 TAAATCAAT 1903
D 3672 TAACTTCA--AAATTTAT 3616
Q 1904 ATCAT 1963
D 3615 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 3556
Q 1964 TATTAAT 2023
D 3555 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 3496
Q 2024 AATTC--TATCATCTGAT 2081
D 3495 AATTCAT 3436
Q 2082 CCAAT 2141
D 3435 CAAT 3376
Q 2142 CAAT 2201
D 3375 TAAATCCAAATCTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 3316
Q 2202 TACTTAAAT 2261
D 3315 CAAT 3256
Q 2262 ATCAAT 2321
D 3255 TACTTAAAT 3197
Q 2322 ACAACAAAT 2381
D 3196 ACCGAAAT 3137
Q 2382 AT 2441
D 3136 GACAAAT 3077
Q 2442 ACAT 2501
D 3076 CTAAT 3017
Q 2502 ATTTAGGCTAT 2515
D 3016 AACTAATAAT 3003

RESULT 27
ACA18468
ID ACA18468 standard; DNA; 1020 BP.

AC AC18468;
DT 19-JUN-2003 (first entry)
DE Prokaryotic essential gene #125.
KW Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.
OS Enterococcus faecalis.
XX MO200277183-A2.
XX 03-OCT-2002.
XX 21-MAR-2002; 2002MO-US009107.
XX 21-MAR-2001; 2001US-00815242.
XX 06-SEP-2001; 2001US-00948993.
XX 25-OCT-2001; 2001US-0342923P.
XX 08-FEB-2002; 2002US-00072851.
XX 06-MAR-2002; 2002US-0362695P.
XX (ELIT-) ELITRA PHARM INC.
XX Wang L, Zamudio C, Malone C, Hasselbeck R, Ohlsen KL, Zyskind JW;
XX Pi Wall D, Trivick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
XX P-PSDB; ABU14598.
XX New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids, required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.
XX Claim 14; SEQ ID NO 6338; 1766bp; English.
XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for
XX proliferation; (7) identifying a compound that influences the activity of
XX the gene product or that has an activity against a biological pathway
XX required for proliferation, or that inhibits cellular proliferation; (8)
XX identifying a gene required for cellular proliferation or the biological
XX pathway in which a proliferation-regulated gene or its gene product lies
XX or a gene on which the test compound that inhibits proliferation of an
XX organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX compound's activity; (11) a culture comprising strains in which the gene
XX product is overexpressed or underexpressed; (12) determining the extent
XX to which each of the strains is present in a culture or collection of
XX strains; or (13) identifying the target of a compound that inhibits the
XX proliferation of an organism. The antisense nucleic acids are useful for
XX identifying proteins or screening for homologous nucleic acids required
XX for cellular proliferation to isolate candidate molecules for rational
XX drug discovery programs, or for screening homologous nucleic acids
XX required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
XX *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
XX prokaryotic essential genes. Note: The sequence data for this patent did
XX not form part of the printed specification, but was obtained in
XX electronic format directly from WHO at
XX ftp.who.int/pub/published_pct_from_who

Best Local Similarity 49.1%; Pred. No. 1.6e-14;
Matches 483; Conservative 0; Mismatches 491; Indels 10; Gaps 3;

QY	629	AAATGAAATTCAGTATTTTAGTGGTCAGAGATCATTTGGCAGACAAATGCAATTCAGTGG	688
DB	4	AAACAAAAGTCGCTGTTTAGTGGTCCAGGTTCCTGGGGGACAGCTTTTGCTCAAGTATTA	63
QY	689	TCAGACATGGATATCATAGTAACTAATGGGGAAGTGATCTAATGAAATTAATACATATA	748
DB	64	GCCGAAATGGACACGAGCTCTGATTTTGGGGAATTAACAGAGCAATGCATGAAATTA	123
QY	749	AACACTTACGAAAAAATTTAAATATTACCCACATATCATCTACGACAAACATATAT	808
DB	124	AATCAAAAACACACCAACAAACATTAATTTACAGATTAATCTTAACCTATCTCAATTGAG	183
QY	809	GCAACCGACCAATATAGACGAGTATTATCTGACAAATAAC-TATGTTATCTTAATCTA	866
DB	184	GCAACCACTGTTTAGCCACTGCTTAGTAATGTGATGCGGATCTTTTGTGTACCA	243
QY	867	TTCTTACCAACAAATTAACGACCATATGTACAAATTAACAACAAACAGATATGTGA	926
DB	244	ACTAAGCCATTCCTTCAGTACCAAGAAATGGCCCACTTTAAAAACAAACCAATT	303
QY	927	AAATATCTCCAAATTAATTTGTGTAAGGATATCGAAATTAATCACTCACTCAATTTCCCA	986
DB	304	ATTATTCATGCAAGTAAAGGCTCGAACAAAGTACACTAAAGCGATTCGAAAGTATA	363
QY	987	GTGAATATGCAAGAAATTTTACAAATTAATCAATTTTATATCTCTGTGCTCAAGTT	1048
DB	364	GCGAAGAGATTCCAGACGA--AAACGGCAAGGATTTGTGTTTATCTGTCCAAAGCC	421
QY	1047	TTGCTAAGAATATGCGAGACATCTCTCTGTAGTATAGTACTTGTGATGATATAAAG	1106
DB	422	ATGCAAGAAAGTGCCTGTTTATGACATTAACACCATTAACAGCAAGCAAGAAATTTAG	481
QY	1107	AACCTGTGAATCATTTGATAGAAAACAATAGTAAATGATGTTCTTAAAAATTAATACATC	1166
DB	482	CAGACGCTGTATGTGCAAGATTTGTTATGATGATTACTTAAATTAACGAATG	541
QY	1167	AAGATTTATATGAGTGCAGATTGAGCTGATTAAGAAACATAATGCAATTCAGATGTG	1222
DB	542	ATGATGTATTTGGTGTGAAACAGGGGCGCTTAAATAATTAATGCTTTGGCGCG	601
QY	1227	GAATTAATCGCTGAAAAAATTTAGTATATAATGCTGTGCTACTGTATTAATCTAAAGCA	1286
DB	602	GAGGATTCATAGCTTAGGCTTTGTGTATATGCAAAAGCGGCAATTAATGACTCGTGTG	661
QY	1287	TGAATGAATTAATAACCTATATATATACCAAAAAATCAATCAATGATGCTTATCATCTTAA	1346
DB	662	TAGCAAGAAATTAATGTCGTTTAGGGGTGGATGG-----AGCAAAACCATTAACATTTA	715
QY	1347	TTGCTCATCATGTCCTTGGAGATCTAAATATTACATGTACAAAGAAATTCACGCAATA	1406
DB	716	TTGCTTTAAGTGTGGGATTTAATGTTACTTATACAGTGCATATCTCGAAATY	775
QY	1407	TGCGTTTGGACTGAATATGGAATAGGTGAATATATAATACCTAATATGATCAACC	1466
DB	776	GCGGTGGGTAATTTATTAGAAATAGCCCATAGTTAGATGAAGTCTTAATAAACATGG	835
QY	1467	TAAAGCTGTGGAAGAACAGTACTGTAAACCACTGATATCATTTAGCAAAAAACCTTA	1526
DB	836	GCAATGTTGTGAAGGTGTTCAACCACTAAAGCAGGCTATGATTTGGCAACAACATAG	895
QY	1527	ATGTAGAATCAACCAATTTGCAATCTATTATTAACAATTAATACAGAAATATATCACTAG	1586
DB	896	AAGTCGAAATGCCAATCAAGAAACAAATCTACATATGCTTGTACATGACGAAGATGAC	955
QY	1587	ATAAAGCCATATCAACATATTAAT	1610
DB	956	AACAAAGCAAAAAGAAATCATGT	979

AAS5118
 ID AAS5118 standard; DNA; 1023 BP.
 AC AAS5118;
 DT 13-FEB-2002 (first entry)
 DE Enterococcus faecalis DNA for cellular proliferation protein #546.
 XX
 KM Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;
 XX antibacterial; drug design.
 XX Enterococcus faecalis.
 OS
 PN W0200170955-A2.
 PD 27-SEP-2001.
 PF 21-MAR-2001; 2001WO-0509180.
 PR 21-MAR-2000; 2000US-0191078P.
 PR 23-MAY-2000; 2000US-0206848P.
 PR 26-MAY-2000; 2000US-0207727P.
 PR 23-OCT-2000; 2000US-0242578P.
 PR 27-NOV-2000; 2000US-0253625P.
 PR 22-DEC-2000; 2000US-0257931P.
 PR 16-FEB-2001; 2001US-0269308P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 PI Haselbeck R, Chisen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 DR WPI; 2001-611495/70.
 XX P-PSDB; AAU35259.
 PT New polynucleotides for the identification and development of
 XX antibiotics, comprise sequences of antisense nucleic acids.
 PS Claim 27; SEQ ID NO 6755; 511bp; English.
 XX
 CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the genes,
 CC their use in the discovery of novel antibiotics, the essential genes,
 CC themselves and the encoded proteins. The prokaryotes used are Escherichia
 CC coli, Streptococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
 CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
 CC useful for the identification of potential new targets for antibiotic
 CC development. The antisense nucleic acids can also be used to identify
 CC proteins used in proliferation, to express these proteins, and to obtain
 CC antibodies capable of binding to the expressed proteins. The proteins can
 CC be used to screen compounds in rational drug discovery programmes. The
 CC antisense nucleic acid sequence is also useful to screen for homologous
 CC nucleic acids which are required for cell proliferation in a wide variety
 CC of organisms. The present sequence encodes an essential prokaryotic
 CC cellular proliferation protein. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pat_sequences
 XX
 SQ Sequence 1023 BP; 344 A; 172 C; 237 G; 270 T; 0 U; 0 Other;
 Query Match 3.1%; Score 148.4; DB 4; Length 1023;
 Best Local Similarity 49.1%; Pred. No. 1.6e-14;
 Matches 483; Conservative 0; Mismatches 491; Indels 10; Gaps 3;
 QY 629 AAAATGAAATCGTATTTAGTGCAGATCATTTGGACAGCAATTCGACTG 688
 DB 4 AAACAAAAGGCTGCTTTAGGTCAGGTCCTGGGGACGCTTTAGCTCAAGTATTA 63
 QY 689 TCAGCAGATGCTATATCACTTAAGTATGAGGACGATCATAGAAATTTACATATA 748
 DB 64 GCGGAAAATGACACAGAGGCTGTATTTGGGAAATTAACGAGCAATCATGATATA 123

QY 749 AACCTTACCGAAAAATTTAAATTTTACCACATATCATCTACAGCAACATATAT 808
 DB 124 AATACAAAACACACCAACAAACATTTATTTACGAAATTTATCTTACTTCAATTCAG 183
 QY 809 GCAACGACATATATAGCAGAGATATATCTGACCAACATAC--ATGATATATCTACTA 866
 DB 184 GCAACGACATATATAGCAGAGATATATCTGACCAACATAC--ATGATATATCTACTA 243
 QY 867 TTCCATACACACATTTAGCAGCATATGTACACAAATACACAAACAGCATATGTGA 926
 DB 244 ACTAAAGCCATTCGTTAGTACAGAGAGTGGCCCAATTTAAACAAACCAATT 303
 QY 927 AAATATCTCAATATTTATTTAGTAAAGATATGAAATTTACATCATCAATTTCCCA 986
 DB 304 ATTATTCATGCAAGTAAAGGCTCGAACAAGGATCAATAGCGATTTACAGAGTATA 363
 QY 987 GTGAATAGCAGAGAAATTTTACAATATATCAATTTTACTCTGTGTCAGATT 1046
 DB 364 GCGGAAAGATTCACAGACA--AAACGCAAGGATGTGTTTATCTGTCAGAGCC 421
 QY 1047 TTGCTAAGAAATTCAGAAACATCTCTTGTATATATCTTGTGCTGTGATATAAG 1106
 DB 422 ATGAGAGAGAGTGTGCTTATGACATTAACCACTACAGCAGACGAAATTTTG 481
 QY 1107 AACTGTGATGATCTTATGAAACAAATATATGATGTCTTAAATATATATACATC 1166
 DB 482 CAGACGCTGTATGTATGCAAGAGTGTATATGATATATTAATTTATACGAATG 541
 QY 1167 AAGATATATAGGTATGACATTTGAGCTGCATTTAAAGAACTATTCATTCATGTC 1226
 DB 542 ATGATGATGTGTATGAAACAGGGCGCCCTTAAATATATTTGCTTGGGCGCG 601
 QY 1227 GAATTAATCGCTGAAAAAATTTAGTAAATATGCTGTCTACTGTATATACATTAAGCA 1286
 DB 602 GAGGATTCATGCTTAGCGTTTGATATATGCAAAACGCGAATTAATGACGTGAGTT 661
 QY 1287 TGAATGAAATTAACACATATATATAGCAAAAAATCATCATATGATCTTCATCATTA 1346
 DB 662 TACAGAAATTAATGCTGTATAGGGGTGCGATGGG-----AGCAACCATTAACATTTA 715
 QY 1347 TTGCTCATATGCTTGTGAGATCTATATTAATCATGTACACAGAACTTCAGCAATA 1406
 DB 716 TTGCTTAAATGCTGTGAGATTTAAATTTATTTACTGTGCAAGTCTTCGAAAT 775
 QY 1407 TGGCTTTGACATGAAATTTGAAAGGATGAAATTTAATATCATTTATGATCAACACC 1466
 DB 776 GGGGTGGGGTAAATTTATTTAGGAAAGGCGCATTAATGATGAAGTCTTAGAAACATGG 835
 QY 1467 TAAAGCTTTGAGGAAACAGATCTGTAACCACTGATCATATAGCAAAAAACTTA 1526
 DB 836 GCATGATTTGAGAGGTTTCAACAATTAAGCAACCATGATGATGGCAACAACTG 895
 QY 1527 ATGTAGAACTCAATTTGATATCTATTTACATTTATACATGGAATATATCTGAG 1586
 DB 896 AAGTCGAAAGCAATACAGAAACATCTACATGCTTTGATCAATGACGAAGATGTAC 955
 QY 1587 ATTAAGCATATCAATATATAT 1610
 DB 956 AACACGACGAAAGAAATCATGT 979
 RESULT 29
 AAS51287
 ID AAS51287 standard; DNA; 1026 BP.
 XX
 XX AAS51287;
 DT 13-FEB-2002 (first entry)
 DE Enterococcus faecalis DNA for cellular proliferation protein #64.
 XX
 KM Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;
 XX

KM antibacterial; drug design.
 XX Enterococcus faecalis.
 OS
 PN NO200170955-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 21-MAR-2001; 2001WO-US009180.
 XX
 PR 21-MAR-2000; 2000US-0191078P.
 PR 22-MAY-2000; 2000US-0206848P.
 PR 26-MAY-2000; 2000US-0207727P.
 PR 23-OCT-2000; 2000US-0242578P.
 PR 27-NOV-2000; 2000US-0253625P.
 PR 22-DEC-2000; 2000US-0257931P.
 PR 16-FEB-2001; 2001US-0269308P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GU;
 PI Yamamoto RT, Xu HH;
 XX
 DR WPI: 2001-611495/70.
 DR P-PSDB; AAU33428.
 XX
 PT New polynucleotides for the identification and development of
 PT antibiotics; comprise sequences of antisense nucleic acids.
 PS
 PS Claim 27; SEQ ID NO 3869; 511pp; English.
 XX
 CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the genes,
 CC their use in the discovery of novel antibiotics, the essential genes
 CC themselves and the encoded proteins. The prokaryotes used are *Escherichia*
 CC coli, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,
 CC *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also
 CC useful for the identification of potential new targets for antibiotic
 CC development. The antisense nucleic acids can also be used to identify
 CC proteins used in proliferation, to express these proteins, and to obtain
 CC antibodies capable of binding to the expressed proteins. The proteins can
 CC be used to screen compounds in rational drug discovery programmes. The
 CC antisense nucleic acid sequence is also useful to screen for homologous
 CC nucleic acids which are required for cell proliferation in a wide variety
 CC of organisms. The present sequence encodes an essential prokaryotic
 CC cellular proliferation protein. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pat_sequences
 CC
 CC Sequence 1026 BP; 346 A; 172 C; 239 G; 269 T; 0 U; 0 Other;
 SQ
 Query Match 3.1%; Score 147.2; DB 4; Length 1026;
 Best Local Similarity 49.1%; Pred. No. 2,4e-14;
 Matches 480; Conservative 0; Mismatches 488; Indels 10; Gaps 3;
 QY 635 AAAATGAGTATTGAGTGCAGATGATTTGGACAGCAATGCAATGCACTGTACACA 694
 DB 4 AAAGTGCAGTGTGTTAGTCCAGGTTCTCGGGGACAGGTTTGAAGTCAAGTATTACCGAA 63
 QY 695 CATGATATCATGTTAACTTATGGGACGCTGATCATAGAAATATACATATAAAGCT 754
 DB 64 AATGGACAGAGGCTGTATTTGGGAAATTAACAGACCAATGATGAAATTAATACA 123
 QY 755 TACCGAAAAATTTAAATATTATCCACATATCATCTACAGACAAATATATGCAAC 814
 DB 124 AAACACACCAACAAACATATTATTCAGATTAATCTTACTTCAATTCAGCAAC 183
 QY 815 AGCAATATGACAGATATTATCTGACAAATATC-ATGTATTATCTTAATATTCCTA 872
 DB 184 ACTGATTAGCCACTGCTTATAGATGTGATGGGATGATCTTTGTGTGATCAACTAAA 243
 QY 873 CACAAACATTAAGCACAATATATGACAAATATCAACAAACAGCATATGTATAAATA 932

DB 244 GCCATTCGTTGAGTACAGCAAGAAAGTGGCCCAACATTTAAAAACAAACCAATTTATT 303
 QY 933 CTCCAATTTATTTAGTATGTAAGGTAATCGAAATATACATCATCTGAAATTTCCAGTAAA 992
 DB 304 CATGCAAGTAAAGGGCTCGAAACAAAGTACACATTAAGCGGATTTCCAGAAATGATCCGAAA 363
 QY 993 TAGCAGAAGAAATTTTCAATATATATCCATTTTATCTCTCGTCCAGTTTCTTA 1052
 DB 364 GAGATTCAGACAGA-AAAACGGCAGAGATTTGGTTTATCTGTCACAAACCATGAC 421
 QY 1053 AAGAAATTCAGAAACATCTCTTGTAGTAATGACTCTGCTGTGATTAATTAAGAACTTG 1112
 DB 422 AAGAAAGTCGCTGTTCAATGACATTCACATTCATACAGACAGAAATTTAGCAGAC 481
 QY 1113 GTGATCATTTGATAGAAACATATAGTAATGATGTTCTTAATAATATATACATCAAGATA 1172
 DB 482 CTGTGATGTGCAAGAGTGTGTTATGATGATTAATTTAGAAATTTATACAAATGATATG 541
 QY 1173 TTATAGGTGTACAGATTTGAGCTGCATTTAAAGAACATTAATTCATATGTAATAA 1232
 DB 542 TGATTTGGTGTAGAAACAGGGGGCGGCTTTAAAAAATATTTATTTGGGGCGGACGCA 601
 QY 1233 TCGCTGCAAAAAATTTAGTATTAATGCTGTGCTACTGTATTATTAAGCATGTAATG 1292
 DB 602 TTCAATGCTTAGGCTTTGTGATTAATGCCAAAGCGCAATTAATGACTGTGTTAGCAG 661
 QY 1293 AAATTTAAACACTATATATATGCAAAAAATATCATATAGATCTTCATATATTAATGCTC 1352
 DB 662 AATTTAGTCGTTTAGGGGTTGCCATGGG-----AGCAACCCATTAACATTTATTTGTT 715
 QY 1353 CATATGCTTTGAGATCTTAATATTAATTAACATGTAACAGAAACATTCACGCAATGCTT 1412
 DB 716 TAACTGTGTGGGATTTTATTTGTTATGTTACATGTCATTCGCAATTTGGGGTG 775
 QY 1413 TTGCATAGAAATTTGAAAGATGAAATATTAATATTAATTAATAGATCACAACTTAAGC 1472
 DB 776 CGGTAATTTATTTAGAAAAAGCCATTAATGATGAAAGCTTTAGAAAAATGCGGCTGA 835
 QY 1473 TTGTGAAGAACCACTGCTGTAATAACCACTGATATTCATTAGCAAAAAAATTAATG 1532
 DB 836 TTGTGAAGAGGTGTTCAACAACATTAAGCAACCTATGATTTGGAACAACAATGAAAGTC 895
 QY 1533 AACTACCAATTTGCATATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1592
 DB 896 AAATGCCAATTCACAGAAACATTTCAATGATCTTTGTAATGACGAAGATGTACACAG 955
 QY 1593 CCATATCAAAACATATTAT 1610
 DB 956 CAGCAAAAGAAATCATGT 973
 RESULT 30
 ABLJ32863/C
 ID ABLJ32863 standard; DNA; 6132 BP.
 XX
 XX ABLJ32863;
 AC
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Human immune system associated gene SEQ ID NO: 836.
 XX
 XX Human; immune system disease; cytosine methylation; antiasthmatic;
 KM antiarteriosclerotic; anti-HIV; anticonvulsant; ophthalmological;
 KM neuroprotection; anti-HIV; anticonvulsant; ophthalmological;
 KM antiarteriosclerotic; anti-HIV; anticonvulsant; ophthalmological;
 KM antiarteriosclerotic; anti-HIV; anticonvulsant; ophthalmological;
 KM antiarteriosclerotic; anti-HIV; anticonvulsant; ophthalmological;
 KM acute myeloid leukemia; Alzheimer's disease; AIDS; epilepsy;
 KM neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
 KM ds.
 XX
 XX Homo sapiens.
 XX

PN WO20020928-A2.
 XX 03-JAN-2002.
 PD
 XX 02-JUL-2001; 2001WO-EP007537.
 PF
 XX 30-JUN-2000; 2000DE-01032529.
 PR 01-SEP-2000; 2000DE-01043826.
 XX
 XX (EPig-1) EPIGENOMICS AG.
 PI Olek A, Piepenbrock C, Berlin K;
 DR WPI, 2002-130909/17.
 XX
 PT Nucleic acid comprising fragment of chemically modified gene, useful for
 PT diagnosis and treatment of diseases associated with abnormal cytosine
 PT methylation.
 PS
 XX Claim 1; SEQ ID NO 836; 32bp + Sequence Listing; German.
 CC The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 CC leukemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 CC diseases. The present sequence is a gene of the invention
 XX
 SQ Sequence 6132 BP; 1642 A; 28 C; 1246 G; 3216 T; 0 U; 0 Other;
 Query Match 3.1%; Score 147.2; DB 6; Length 6132;
 Best local Similarity 44.2%; Pred. No. 2.6e-14;
 Matches 1069; Conservative 0; Mismatches 1318; Indels 31; Gaps 10;
 QY 281 AAGAAACAAGTATGACAAAGATACCTCTCTCTGTCATCCATATTTCCATTTAT 340
 Db 5037 AACAAACCTTTATCTTCAACACCTCTTAATTTTCTCCCAATCCCTAAATAAT 4978
 QY 341 TTACAAATATCATAGAGTTCTTTATTCGACATCATTTAAATCAATCAAAACAAATA 400
 Db 4977 ATTAATAATTAATTAATA--TTTATTAACACTTAATTAATTAATAATAATAATA 4921
 QY 401 TAGTACTATCTATTATATAAAAAATCCTCAATAGCTCTTAGCAAGAACTCAATATA 460
 Db 4920 AAAAAGTTACTCAACCTTAATAAAAAATTAATAAACTATAAAAAATAGTACTATCAT 4861
 QY 461 TGAAGACCTTAATCTATTATTAATTTTCAAAATTTTATCTACTCATCATATTAGCAT 520
 Db 4860 TTATATACAC-AAACAATTAATTAATCTTAATTAATAAACAATTAATTAATAAACA 4802
 QY 521 AATATCACTGTCGACATGATTAACAATAGTACATATTTCTTAACATAGATTAATATCA 580
 Db 4801 ACTACCAAAACAACTCAAAAAATTAATAAAATCCAAACAAATCTCACTCATCTA 4742
 QY 581 ATTAATATATTAATTTTGAAGTCTTCTGATTAATTAATAGTAAAAATGAAAAATC 640
 Db 4741 CTATTCATATATACAC--TAAATTCCTTAACAAACAAATTAATAAAAAATTAATA 4684
 QY 641 AGAATTTAGTGCAGATCATTTGGGACAGCAATAGCAATGCACTGTCGACATAGGT 700
 Db 4683 AAAATATCAATTAATTAATAAAAAATTAATACTATTTCTATTTCATTAATTAATCAT 4624
 QY 701 ATATCAATTAATTAATGGGAGCTGATCATAGAAATATTAACATTAATAAACAATTA 760
 Db 4623 ATACCAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4564
 QY 761 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 820
 Db 4563 CAAATTTCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4504
 QY 821 ATAGACGAAGTATTAATGACAAACATATAGTATTAATTAATTAATTAATTAATTA 880

Db 4503 AAAAAAATAATAA-----AAACAATTCCTTTATTAATTAATCCCAAAAAATTAATAATC 4448
 QY 881 TTACGACCATATGTAACAAATTAACAACAAAGCATATGTTAAATTAATCTCAATA 940
 Db 4447 ATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4388
 QY 941 TTAATTTAGTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 1000
 Db 4387 CTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4328
 QY 1001 GAATTTTACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1060
 Db 4327 CAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4278
 QY 1061 GCAGAAATCTTCTCTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1120
 Db 4277 TCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4218
 QY 1121 TTGATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1180
 Db 4217 AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4158
 QY 1181 GTACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1239
 Db 4157 TCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4098
 QY 1240 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1299
 Db 4097 AAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4038
 QY 1300 AACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1359
 Db 4037 TTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 13978
 QY 1360 TCTTGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1419
 Db 3977 TCTTCAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3918
 QY 1420 AGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1477
 Db 3917 CTACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3858
 QY 1478 GAAAGACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1537
 Db 3857 AAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3798
 QY 1538 CCAATTTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1597
 Db 3797 AATTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3738
 QY 1598 TCAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1657
 Db 3737 ACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3678
 QY 1658 GTCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1717
 Db 3677 TACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3618
 QY 1718 TAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1777
 Db 3617 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3558
 QY 1778 CTAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1837
 Db 3557 CAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3498
 QY 1838 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1897
 Db 3497 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3438
 QY 1898 CACTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1957

Db 64774 --ATCCACAACTTTTACGGGCTAAGTGTGATGATCGTTACTTGTACAA 64718
 QY 1389 CAGAACATTCACGCAATATGCTTTGGACTGAAATAGAAAAAGGTGAATATAA 1448
 Db 64717 GGTGCGATTCACTGATTTGGCGTGTGAAACATCTCGTTAAAGCGAATTTAGATG 64658
 QY 1449 CATTAAATAGATCAACAACCTTAAAGCTTTGTAAGAACACGATCTGTAATAT 1508
 Db 64657 AAGTTTAAAGAAAAATGGGTATGCTGTGAAGGTGTCGTACAGCCAAAGCAGTTGATG 64598
 QY 1509 CATTAGCAAAAAAATCTTAATAGAACTACCAATTTGATCTTATTAATTAATAC 1568
 Db 64597 GTTGGCGGAAAAAATAGATATTTGATATGCCAATTAACGAAATCGATTTTAT 64538
 QY 1569 ATGAGATATTA 1579
 Db 64537 TCGAGAAATATA 64527
 RESULT 32
 AB067194
 ID AB067194 standard; DNA, 319630 BP.
 AC AB067194;
 XX
 DT 29-AUG-2002 (first entry)
 XX
 DE *Listeria innocua* contig DNA sequence #7.
 XX
 KM Antibacterial; *Listeria*; food contamination; mutational analysis;
 KM infection; ds.
 XX
 OS *Listeria innocua*.
 XX
 PN WO200228891-A2.
 XX
 PD 11-APR-2002.
 XX
 PF 04-OCT-2001; 2001WO-FR003061.
 XX
 PR 04-OCT-2000; 2000FR-00012697.
 XX
 PA (INSP) INST PASTEUR.
 XX
 PA (CNRS) CNRS CENT NAT RECH SCI.
 XX
 PI Kunec F, Glaeser P;
 XX
 DR WPI: 2002-332479/37.
 XX
 PT New genomic sequences from *Listeria* species, useful for detection,
 PT treatment and prevention of infection, also related polypeptides,
 PT antibodies and modulators.
 XX
 PS Claim 5; SEQ ID NO 7; 180bp; French.
 XX
 CC The present invention relates to nucleic acid sequences (AB067188-
 CC AB071222) from *Listeria* sp. The sequences are useful as probes and
 CC primers for identification and/or detection of *Listeria* (e.g. as
 CC contaminants in foods, or mutational analysis) and for analysis of gene
 CC expression. Proteins encoded by the nucleic acid sequences can be used to
 CC screen for compounds that modulate gene expression, replication and
 CC pathogenicity of *Listeria* (potential therapeutic agents), also for
 CC treating infections by *Listeria*, and are useful as immunogens in anti-
 CC *Listeria* vaccines. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at http://wipo.int/publ/published_pcf_sequences
 CC
 SQ Sequence 319630 BP; 105207 A; 55428 C; 66726 G; 92263 T; 0 U; 6 Other;

Query Match 3.1%; Score 146.2; DB 6; Length 319630;
 Best Local Similarity 49.8%; Pred. No. 4.5e-14;
 Matches 484; Conservative 0; Mismatches 473; Indels 14; Gaps 4;

QY 613 GATATTAATAGTAAATAATGAAATATCAGTATTTTATGCTGAGATCATTTGGCACAGC 672
 Db 51941 GAAAAATATATATGACACAGAAAAAGTGTCTATTTCTGGCGGTGGAAGTGGGAAACAG 52000
 QY 673 AATAGCAATTTGCACTGTGACACATGTATATACATTAATCTATGGGACGATGATG 732
 Db 52001 ACTGCGACTTGTCTGCTGATATATATATTAACCAAGTATTTGGGAACTTAGATA 52060
 QY 733 AATATTAACATATATTAACATTTACCGAAAAATTTAAATATTTAACCAATATCATCT 792
 Db 52061 AATTTGAATGAATTAATGAATTCGACAGATATGCTACTTTGGCAGATATTAATTT 52120
 QY 793 ACCAGCAACATATATATGCAACGACATATTAACGAAATTTATCTGACACAAATACAT 852
 Db 52121 ACCACTGAGTAAAGCACTTATCACTTATATAGCTATATAGATGATGCTGAAT-- 52177
 QY 853 TATTATCTTAATCTATCTTACACAAATTTAGCAGCATATGTACAAATATCAACACNA 912
 Db 52178 TGTGCTGATTTGCTATTTCCAAATATGCAATGGCTGTGTTGTAGCAGCTAAATGAAGC 52237
 QY 913 ACAGCAATATGTAATTAATATCTCAATATTAATTTGTAGTAAAGTATGCAATATCATC 972
 Db 52238 GCTGAA-AGAACCACTATTTATAGTCAATGTAAGGAGTTGAACCAAGAAACAAATC 52296
 QY 973 ACTCAATTTCCAG---TGAATAGCAGAAATTTTACATATTAATCAATTTTCA 1028
 Db 52297 TTGAAATGTCAAGATATTAAGAAAGTATGATGCTACAAAGTAAAGCTTTGTTG 52356
 QY 1029 TACTCTGTGTCGAATTTTGTCTAAAGAAATTCAGAAATCTTCTGTGTATATGATC 1088
 Db 52357 TTCTTTCAGGGGCTGATGCTGCAAGTAAAGTGTCTTCTGTATCCAAAGCCTTTG 52416
 QY 1089 TTGCTGTATATATAAGAACTGTGGAATCATTTGATGAACCAATTAATATGATGCTC 1148
 Db 52417 CTAGCTGTAAAGATTTGAAAGCAGCAAGATGTGCAAGATGCTTTATTAATTAATTT 52476
 QY 1149 TAAATTAATATATCAATCAATCAATATTAATGATGATGATGAGCTGCTTAAGAAC 1208
 Db 52477 TACGTATTTATGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 52536
 QY 1209 TAATGCAATTCAGTGAATATATGCTGGAAGAAATTTAGTATATATGCTGTTGCTA 1268
 Db 52537 TTATCGCACTTGGCGCAGAAATCTGTATGCTGCTGATACGCGATATATGCTTAAGCG 52596
 QY 1269 CTGTTATTAATTAAGCATGATGATTAATTAATTAATTAATTAATTAATTAATTA 1328
 Db 52597 CACTATATGACAGCGGAATGCTGAATTAATCTGCTGGAATGCTGCTGCTGCTGCTA-- 52653
 QY 1329 TAGATCTTATATCATTAATTTGCTCATCATGCTTTGGAGATCTATATTAATATGATCA 1388
 Db 52654 ---ATCCACAAACCTTTTACGGGCTACGATGATGATGATGATGATGATGATGATG 52710
 QY 1389 CAGAACATTCACGCAATATGCTTTTGGACTGAAATAGGAAAGATGAAATATTAATA 1448
 Db 52711 GGTGCAATTCAGTATATGCGTGTGCTGCTGGAACATCTCGTAAAGCGGAATTTAGAT 52770
 QY 1449 CATTAAATAGATCAACAACCTTAAAGCTTTGTAAGAACACATCTGTAATTAATCAATG 1508
 Db 52771 AAGTTTAAAGAAAAATGGGTATGCTGTGAAGGTGTGTCACCAAGAGAGTTCATG 52830
 QY 1509 CATTAGCAAAAAAATCTTAATAGAACTACCAATTTGATCTTATTAATTAATTTATAC 1568
 Db 52831 GTTGGCGGAAAAAATAGATATTTGATATGCCAATTAACGAATCGATTTTACGATTTAT 52890
 QY 1569 ATGAGATATTA 1579
 Db 52891 TCGAGAAATATA 52901

RESULT 33
 ABK28332/C
 ID ABK28332 standard; DNA, 11745 BP.

AC ABK28332;
XX
DT 23-APR-2002 (first entry)
XX
DE DNA transcription associated complementary genomic DNA #103.
KM
KM DNA transcription associated gene; peptide nucleic acid; PNA-oligomer;
KM PNA; cytosine methylation state; SND; retroviral infection; gene; ds;
KM single nucleotide polymorphism; adenosine deaminase deficiency; cancer;
KM viral infection; Searcy syndrome; haematological disorder; tuberculosis;
KM immunological disorder; Werner syndrome; developmental disorder;
KM psoriasis; Rieger's syndrome; neurological disorder; erythropoiesis;
KM neurodegenerative disorder; Waardenburg syndrome; Niemann-Pick disease;
KM myelodysplastic syndrome; myocardial infarction; hypertension; arthritis
KM angiodysplasia; congenital heart disease; HDR syndrome; gene therapy;
KM polyglutamine disorder; solid tumour.
XX
XX unidentified.
XX
XX
XX WO200192565-A2.
XX
XX
XX 06-DEC-2001.
XX
XX
XX 06-APR-2001; 2001WO-EP003973.
XX
XX 06-APR-2000; 2000DE-01019058.
XX 07-APR-2000; 2000DE-01019173.
XX 30-JUN-2000; 2000DE-01032529.
XX 01-SEP-2000; 2000DE-01043626.
XX
XX (EPIG-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
XX
XX WPI; 2002-090046/12.
XX
XX
XX New nucleic acids or oligomers, useful for diagnosing or treating
XX diseases associated with DNA transcription, e.g. immunological disorders,
XX Werner syndrome, psoriasis, myocardial infarction, solid tumors or
XX cancer.
XX
XX
XX Claim 1; SEQ ID NO 206; 32pp; English.
XX
XX
XX The invention relates to a nucleic acid, which comprises a segment of chem
XX chemically pretreated DNA of genes associated with DNA transcription from
XX one of 346 sequences, and an oligomer, in particular an oligonucleotide
XX or peptide nucleic acid (PNA)-oligomer that hybridises to or is identical
XX to the chemically pretreated DNA of genes associated with DNA
XX transcription. The set of oligomer probes are useful for detecting the
XX cytosine methylation state and/or single nucleotide polymorphisms (SNPs)
XX in a chemically pretreated genomic DNA. The nucleic acids are useful for
XX diagnosing or treating diseases associated with DNA transcription
XX (particularly with the methylation status), e.g. adenosine deaminase
XX deficiency, viral infection, retroviral infection, Searcy syndrome,
XX haematological disorders, immunological disorders, Werner syndrome,
XX tuberculosis, developmental disorders, psoriasis, Rieger's syndrome,
XX neurological disorders, neurodegenerative disorders, Waardenburg
XX syndrome, Niemann-Pick disease, myelodysplastic syndrome, myocardial
XX infarction, hypertension, angiodysplasia, erythropoiesis, congenital heart
XX disease, HDR syndrome, arthritis, polyglutamine disorders, solid tumours
XX or cancer. Sequences ABK28127-ABK28472 represent DNA transcription
XX associated genetic DNA molecules of the invention. Note: The sequence
XX data for this patent did not form part of the printed specification but
XX was obtained in electronic format directly from the European Patent
XX Office
XX
XX Sequence 11745 BP; 3183 A; 270 C; 2383 G; 5908 T; 0 U; 1 Other;
SQ
XX
XX Query Match 3.1%; Score 145.4; DB 6; Length 11745;
XX Best Local Similarity 43.5%; Pred. No. 5,2e-14;
XX Matches 966; Conservative 0; Mismatches 1241; Indels 18; Gaps 6
XX
XX 328 ATTTCGAATATTATTAACAATATCATGAGCTCTTATTTCGACATPACATATTAAATCAAT 387

[illegible]

[illegible]

QY	2545	AGTTCGT	2551	
Db	3824	AATTCCT	3818	
				RESULT 34
				ABL33013/C
				ID ABL33013 standard; DNA; 7597 BP.
XX				ABL33013;
XX				26-MAR-2002 (first entry)
DE				Human immune system associated gene SEQ ID NO: 986.
XX				Human, immune system disease; cytosine methylation; antiasthmatic;
KM				antiarteriosclerotic; antianaemic; cytosolic; noctropic;
KM				neutroprotective; anti-HIV; anticonvulsant; ophthalmological;
KM				antiheumatic; antiarthritic; antidiabetic; antipsoriatic;
KM				antiflammaroy; cancer; eye disease; arteriosclerosis; anaemia;
KM				acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KM				neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
XX				ds.
OS				Homo sapiens.
XX				WO200200928-A2.
XX				03-JAN-2002.
XX				02-JUL-2001; 2001WO-EP007537.
PF				30-JUN-2000; 2000DE-01032529.
XX				PR 01-SEP-2000; 2000DE-01043826.
XX				(EPIG-) EPIGENOMICS AG.
PA				Olek A. Piepenbrock C, Berlin K;
XX				PI 2002-130909/17.
XX				DR Nucleic acid comprising fragment of chemically modified gene, useful for
XX				diagnosis and treatment of diseases associated with abnormal cytosine
PT				methylation.
XX				Claim 1; SEQ ID NO 986; 32bp + Sequence Listing; German.
PS				The present invention provides a number of human immune system associated
XX				genes which are modified by the methylation of cytosines. The sequences
CC				can be used in the diagnosis and treatment of immune system disorders,
CC				including eye diseases such as retinopathy, neovascular glaucoma and
CC				macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC				leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC				rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC				diseases. The present sequence is a gene of the invention
XX				Sequence 7597 BP; 1965 A; 109 C; 1790 G; 3733 T; 0 U; 0 Other;
SQ				
				Query Match 3.0%; Score 144.2; DB 6; Length 7597;
				Best Local Similarity 47.4%; Pred. No. 7.9e-14;
				Matches 556; Conservativity 0; Mismatches 613; Indels 4; Gaps 4#;
QY	1294	AATTAAACACTATATATAGCAAAAAATCATTCATAGATCTTCATACATTAAATGGTCC	1353	
Db	1272	AAACAAATATATAAT	1213	
QY	1354	ATCATGCTTGGAGATCTAATATTTAATGATGACACAGAACATTCACGCATATGGCTTT	1413	
Db	1212	ATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATGAT	1153	
QY	1414	TGGACTGAATATGGAAGAAGGTGGAATATATTAATACATTATATAGATCACAACCTAAAGCT	1473	

[illegible]

ID	ABK31511	standard; DNA; 47108 BP.
AC	ABK31511,	
XX		
DT	23-APR-2002	(first entry)
XX		
DE	Signal transduction associated gene modified complementary DNA #177.	
XX		
KM	Human; signal transduction associated gene; cytosine methylation state;	
KW	Cpg island; signal transduction associated disease; solid tumour; cancer;	
KX	antitumour; cytostatic; mutant; ds.	
XX		
OS	Homo sapiens.	
XX	Synthetic.	
PN	WO200200926-A2.	
XX		
PD	03-JAN-2002.	
XX		
PF	29-JUN-2001; 2001WO-EP007472.	
XX		
PR	30-JUN-2000; 2000DE-01032529.	
XX		
PR	01-SEP-2000; 2000DE-01043826.	
XX		
PA	(EPIG-) EPIGENOMICS AG.	
XX		
P1	Olek A, Piepenbrock C, Berlin K;	
XX		
DR	WPI; 2002-147896/19.	
XX		
PT	Oligonucleotide for diagnosis and therapy of diseases associated with	
XX	signal transduction e.g. cancer; comprises chemically modified genomic	
PT	sequences of genes associated with signal transduction.	
XX		
PS	Claim 1; SEQ ID NO 354; 24pp; English.	
XX		
CC	The present invention relates to chemically modified DNA sequences of	
CC	signal transduction associated genes. The DNA sequences are chemically	
CC	modified using a solution of bisulphite, hydrogen sulphite or disulphite.	
CC	Also disclosed are oligonucleotides and/or PNA oligomers for detecting	
CC	the cytosine methylation state (Cpg islands) of these genes, and a method	
CC	for the diagnosis and/or therapy of genetic and epigenetic parameters of	
CC	genes associated with signal transduction. The genomic DNA can be	
CC	obtained from cells or cellular components which contain DNA, e.g. cell	
CC	lines, biopsies, blood, sputum, stool, urine, cerebral-spinal fluid,	
CC	tissue embedded in paraffin such as tissue from eyes, intestine, kidney,	
CC	brain, heart, prostate, lung, breast or liver, histologic object slides,	
CC	and all their possible combinations. The sequences of the invention are	
CC	useful for the diagnosis and therapy of diseases associated with signal	
CC	transduction e.g. solid tumours and cancer. ABK3158-ABK31545 represent	
CC	chemically pretreated genomic DNA sequences of different genes associated	
CC	with signal transduction, or their complementary sequences. Note: The	
CC	sequence data for this patent did not form part of the printed	
CC	specification, but was obtained in electronic format directly from the	
XX	European Patent Office	
XX		
SO	Sequence 47108 BP; 15349 A; 429 C; 8736 G; 22594 T; 0 U; 0 Other;	
XX		
Query Match	3.0%; Score 143; DB 6; Length 47108;	
Best Local Similarity	43.6%; Pred. No. 1.3e-13;	
Matches 1181; Conservative	0; Mismatches 1495; Indels 31; Gaps 11.	
QY	362 TATTCGACATACATATTAAATCAATCAAAAACAATAATAGTACTATCTAATTAATMAA	421
DB	9596 TTTTATATATTAAATATATAATATATAAAAAAATTTAAATCAAAAATTTACTATT	9533
QY	422 AAAATCCTCAATAGTCTTTAGCAAGAACTACAAATATATATAGACACTTAACTATTAT	481
DB	9536 TAAATTCGAATCCACACACTTTCTTAAATTAATCTTAAACAATATATTATTATCTATTACT	9477
QY	482 ATATTTTCAATTTATTTACTACTGCATCATATTAGATATAATACATCGACATAGATT	541
DB	9476 ATATTAACAAACCAATTCCAAATTTTAATTTACTCAAAAACAATATTTTATATTTCTCAT	9417

QY	54	AAQATAGTAAACATTTCTTTAAAGCTGATATATATGCAATTAATATATATATTTTG	601
Db	9416	CTTAATATTAATTAATCTAACTCAACCCAAAATTTCTTATTTCAATATTAACATTTTATTA	935
QY	602	ACTGGTTTCTTGATATATAATAGTATTAATAATGAAATCAGTATTTTGGTGCAGATCA	661
Db	9356	CCCACATTTTATCTTAAATTAATAAAAAACAATACAAATATCTATTTTAAACATTTTCT	929
QY	662	TTTGGACACAGCATTTAGCAATTTGCACTGTCAGACATGTAATTCAGTTAACTTATGGGA	721
Db	9296	TTAAATTAATTTTACTCTCACTACCTTAATATCCCAACCTTCGAAAAACCGAAACAAATAA	923
QY	722	CGATGATCATGAATATTTACATATTAACACTTACGGAATAATTTAAATTTATATCC	781
Db	9236	TCATTAATAATCAAAAATATGAAACCAATCTTAACCAATATATTAATAACCCC-ATCTCACTA	9178
QY	782	ACATATCATATTCACAGACAACTATATGCAACAGCAATATAGACGAAGTATATCTGC	841
Db	9177	AAATATCAAAAAATTAATCTAATTAATTAATPAACAACCTTAATATCCACTACTTAAAA	9116
QY	842	AACATATACATGATATTAATCTTAACTATCTTCCACAGAACTTAGCAGCAATATGTACAA	901
Db	9117	ACTTAATAACAAAAATTAATTAATCCAAAAAGAAAAATTGAATTAACCGAAATATTA	9058
QY	902	ATACAAACAAACAGATATGTGTAAATAATCTCCATATTAATTTGTATGAAGTATC	961
Db	9057	CCATTAACCTCCCACTTAACAAAAAGAAAAAGAACTCCATCTCAAAAAAATTAATAAT	8998
QY	962	GAAATATACATCACTCAAAATTTTCCAGTAAATATAGCAAGAAATTTTACAA-TATATCC	1022
Db	8997	CATTAATACCTAATCACTATTAATAACTCAATTAATTAATTTTTTTAAATTTATATTA	8938
QY	1021	AATTTTATTAATCTCTGCTGCCAAGTTTGGCTAAAGAAATTCGACACATCTTCTGTG	1088
Db	8937	TAAACACATTAATCCCAATACAAATTAATCTATTAATAAATCCATTAATTAATCTTAATTC	8878
QY	1081	TATAGTACTGTGCTGATATTAATAAGAACTTGGAATCAATGATAGAAACAATAAGTA	1140
Db	8877	AACCTTAATTAATTAATTTAAACAATTTTCTTATTTTTCGTAATTAATAATATATAC	8818
QY	1141	TGATGTTCTAAATAATATATACATCAAGATATATATGGGTACAGATTGGAGCTGAT	1200
Db	8817	TATTAATCTTTTATTAATTAATCAATTCATTAACAAATTAATTAATAATTTTATATTTAA	8758
QY	1201	AAAGAACATATATGCAATTTGCAATGGAATATCGCTGGAAAAAATTTAGGTATATATGC	1266
Db	8757	AAATATTAATCTTAAAAAATTAATTTCTTACCCCAATAAATTTATTTCTTAATTAATA	8698
QY	1261	TGTTGCTA---CTGTATTAATCTAAAGCAATGAATGAATTAATAACCTATATATGCA	1318
Db	8697	CAAAAAAATAATTAATAATTAATAACCAATATTTTCAATTAATAACCAATATCTATTAATA	8638
QY	1317	AAATCATCTCAATAGATCTTTCATACA-----TTAATTTGGTCATCATGCTTGGAGAT	1368
Db	8637	ACATCAAAACAAAATTAATAAAAAAACAATACATATATTAATTAATAAATTTACTCTT	8578
QY	1370	CTAATTTTAACATGTACAAAGAACATTCAGCAATATGCGTTTGGACTGAGAAATAGGA	1422
Db	8577	CTATTTTAATAATATCAAAAAAACCCTTTCTATATATTAATCAATTAACATTAACAA	8518
QY	1430	AAAGGTAGAAATTAATATCATTTATATAGATACACAACCTAAAGCTGTGTAAGAACAGAT	1488
Db	8517	ATTAATAAAAAAATAATTAATCTTATATCTCTCTACCTAAACCTCTCT---TCTCCAAA	8468
QY	1490	ACTGTAAAAACCACTGATATCATTAAGCAAAAAAATCTTAATGTGAATCTCAATTTGCAT	1548
Db	8460	TATTTTAAATATCAAAAAATCTTAATAACAAAAACCTATATTTTAATAACAAATTAATAAATCTAA	8400
QY	1550	TCTATTTTCAATTTATATACATAGAAATATATCACTAGTAAAGCAATCAATCAATATTA	1608
Db	8400	TATTAATTAACCAAAACA-----AACAAATTAATAAATTAATAAATTTTCAATATTA	8348

QY	1610	TCTTAGTGTATCATCCTTGATTAATTTCAACACATGTAATTAACGGTCAAAAATATA	166
Dp	8346	CTAAAACCTACCCCTTAAACCATTATTAATAAATTAATAAATACTACAAAATTTT	828
QY	1610	GAATACAAAGCTTCAACCACATATATAAAAAGACGTCAAAATCCATCGTTAACATCAACT	172
Dp	8286	AAAAAATAAATACATAATCTAAATTTATATATTAATTAATTCATCAACCTACTAAATTT	822
QY	1730	AAGTAGTGAACATATTTCATACACAACTTAAGTATTAATTAAATTTCCCT-AGCTAATTA	178
Dp	8226	AAAAATTAATACAAAAATTAATAAACAAAAACAAAAAAACCTCTTATATATA	816
QY	1789	TATTAACCAATTCCTATCTCCTATCTCATTTAACCAATTGTTATCAAGTTATAAATAGT	184
Dp	8166	CAAAATAAAAAATAAAAAATTAATTAATTAATAATAAATAAATAAATAATATACATTTATA	810
QY	1849	CAAAATTAACCTTTACTATTTATATAAA--TAATATGTTTGACACCTTTTCACTACATC	190
Dp	8106	AAAATTAACCTTTACTATATAATAAAAAACCACTAAATTTTAAAAATAATAATACATA	804
QY	1907	ATGTAATTAAGACATATATTTCATATAAAATACTATATAAAGAACCATACATATAT	196
Dp	8046	ATCTAATTTATATATCAAAATTAATATCACTCAACCTACPAATTAATAAATAATTAATAA	798
QY	1967	TAAATTAACCAACCTATTTACTAGTATATATAGATGTCCTGACATATAAT	202
Dp	7986	ATAAAACAAAAACAAAAATTAATTAATAAATTTCTTCAATATCCAAATAAAAATATA	792
QY	2027	TCTATCATCCTGACATTTTGCAAACATTAATATGATTAACGTTTTTTATGCCAAT	208
Dp	7926	ATATATTTAAATCAAAATTAATACATTAATAATTAATTCATTTCTATATATATACACAT	786
QY	2087	AGTTCTTAATTAATTTCTATTTACCCAAAAAGTTAAATACATTTCTCGACACATA	214
Dp	7866	A-TTTTAAAAATAAATCAATTAATTAATTTCTAATATATAAATAAATAAATAAATAATTA	780
QY	2147	TGAAGAAGATACATACACGTCAATTAATTAACAATTTCTAATATTAACCACTCTC	220
Dp	7807	AATTTAATTTCTCACAAATTAATAACCTTAATAAATAAATAAATAAATAATATCGTTTATA	774
QY	2207	TAAACATAGAACCAATAGTATATCATACCTAGTAAGATATCCCTACAGCATATACA	226
Dp	7747	ACATTAATTAACATATATAAATAAATAATTAATTAATAAATAAATAAATTTTCTTAACAAT	768
QY	2267	ATATTAAGAAAAATCGTTATGTTATTTATTTATACATTAACAATTCCTCACTACAAC	232
Dp	7687	TCAATTTAAATACCTATATTAACATTCAAAAAATACTTAACAAACAATTAATTAATTCAC	762
QY	2327	AAAACTCAAGTCTATATTAACCAATTCAGAAATTAATATATCTTCAACAACAATAGT	238
Dp	7627	AAATCTAAAAATTAATAAATAAATAAATCTAATTTAAAAATATATATATATATTAATAA	756
QY	2387	TATCAATTTCTATATATCATCATATCTCTTTTAAAAAGATGACTCAAAATCTATACATA	244
Dp	7567	TAAAAATACCTATTTAAAAACATTAACAATAATTAATCATATAAAAAATTA-ATACA	750
QY	2447	AATATATTAACATGCAATGTTTATTCAGAAACATTAACACATTCAGAAATTTCAATTTA	250
Dp	7508	CATTAATAAATAAATAATTAATTCACAAAACTTAACCTTCAATATCCCAAAATTTAAACGTAA	744
QY	2507	CGGTATAGTGTGAAGCTTCATTCGAAACAAATTCACAGATCGTGATACCTACCTCA	256
Dp	7448	AAAAATTAATAAGAACTTAATTAACAATCTTAATAAATAATATATCAATTAATAAATAA	738
QY	2567	AAACGAGATATCTGAGTTTAAATATCTGCTTTCAATTTCTTCAAGACTTAACGTAT	262
Dp	7388	AAACGAAAAATTTCTAATACCTCAAAATTCAAATTAATTAATTAATTAATTAATTAATTT	732
QY	2627	GGTACAAAAATTAATTTAGACCTTAATAACATCAATTAAGCGAATAGATATCATATCT	268
Dp	7328	AAATATCTACAAATCAATTAATTTAAATAAGTAACGATACATATATTTAATTAACCTTAAT	726
QY	2687	ACTGTAATATCTTTGCTGTATCATCTCCTTAGAAGATTTCAATATTAATTAATCAATCT	274

Db	1970	TTATTAACCAACTATTAACATTCCTCAATATATATTTTCCAAACATACCTTAATTAATAAAATT	1911
Qy	1642	CAATGATTAATTAATAGCGTCAAAATAATATGAATATACAGCGTTACCGCATATATTAATAAGCA	1701
Db	1910	CATATCTTTAATTAACATTAAAAAATCTTATATATATTAATTAATTTTTCACATATTTAAAAAATA	1851
Qy	1702	GTCAAAAT-ACCAGTGTAAACAATCCAACTAGATATGTAACATATTAACATAACAACTTAA	1766
Db	1850	TAAAAATTAACAATTTCTACATCTCCAAAATCTTAAATATATACATTTAGCTTAATTA	1791
Qy	1761	GATTAACCTTAATATTTCCAGTATTTATATTAATAAAGATTCATCTCTCTATTCATTA	1820
Db	1790	ACAAAATCAGAAATCAATCATCATCTTAATATCCAAACCTCTTAATTCGAAATCTTAACA	1731
Qy	1821	ACCAATGTGTATCAGATTAATAATAATAGTCAAAATTAAGTTTACTATTATTAATAATAT	1880
Db	1730	ATTTATTTTAATCTTAACAAATTTTCATTTTCTAATATATATTTTAAATTAACATATC	1671
Qy	1881	ATGTGTGACACTTTTTCACATCATCATGATATATATATGACATATTTTACHTAATAATA	1940
Db	1670	AAATTAATACCTTTTATATATATTAATACATATATCTTATATTTCTTTTCAATTAACATTA	1611
Qy	1941	CTAATTAATAAAGAACCAATTAATATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	2000
Db	1610	ATACACATATACCCCAAAATACCCCAATTTTCTACTACATTAACAAATATTAATAG-AACT	1552
Qy	2001	TACATAGATATGCTCCGACATTAATAATTTCTATCATCTCTGACATATTTCAACATTAATA	2066
Db	1551	ATATTAATTAATCCAAATCTCATTTTACTACATTAACAACAAATATTAATTAATCTTAATA	1492
Qy	2061	TGATTAAGTGTTTTATATATGCCCATTAAGTCTTAATAATAATTTGCTAATTAACCAAAAAAG	2120
Db	1491	AATAAAAAATACCGTAAAAAATATCTTTAATTAATCAATATCAATATTTAATTAATTAAT	1432
Qy	2121	TTAAATTAACACTTCCCTGCAACAATATGAAGAATACATTAACGCTCAATTAATTAATA	2180
Db	1431	TATATTAATCAATTAATTAACAACTTTTAATAAAAAAATTAATTAATAATTAATTAATTAATA	1372
Qy	2181	CAAAATTTCTAATATATACACACTTACTTAACATATAGAAACAAATGATATATCAT-----	2233
Db	1371	AAATTAATTAATTTTCTCAATCTTAATTAATTAATAAATTAATAAACCAGAAATAAATATATCTTA	1312
Qy	2234	ACCTTAAGTAAGTATATCCCTACACAGATATGATCAATAATAAGAAATAATCGTATAGTATAT	2293
Db	1311	ACCTTAATTAATCAATATTAATTTATTTTTCAAAAATTAATTAACACTCTCCCTTAATTAATA	1252
Qy	2294	TTATTAACAATTAACAAAATTTCCAACTTCAACAACAAAACTCAAGAGTCTATTTATTAACAAT	2353
Db	1251	AATTAATAATTTTACCTATTAATAAATACTATTAACATTTTATTAATTAATTAATTAATTAATTA	1192
Qy	2354	TCAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	2413
Db	1191	AAAAATCAATTAACCTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	1132
Qy	2414	TTTTTAATAAAGATGACTTCACAATTCATATCAATTAATAATTAATTAATTAATTAATTAATTA	2473
Db	1131	ATCTCAATTAATTTAATTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	1072
Qy	2474	GAAACATTAACAAATCAAGAAATTTTCAAT	2504
Db	1071	ATTAATTAACATTAATTAATTAATTAATTAATTT	1041
RESULT 37			
ADCG1772			
ADCG1772 standard; DNA; 1062 BP.			
ADCG1772;			
01-JAN-2004 (first entry)			
E. faecium DNA sequence SEQ ID 1399.			

XX ds: Gene; urinary tract infection; bacteraemia; endocarditis; wound,
KW abdominal-pelvic infection.

OS Enterococcus faecium.
PN US6583275-B1.
XX
XX 24-JUN-2003.
XX
XX 30-JUN-1998; 98US-00107532;
XX
XX 02-JUL-1997; 97US-0051571P.
PR 14-MAY-1998; 98US-0085598P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
PI Doucette-Stamm LA, Bush D;
DR WPI: 2003-799836/75.
DR P-P5DB; ADC95426.
XX

PT New isolated nucleic acid derived from Enterococcus faecium encoding an
PT Enterococcus faecium polypeptide useful for detection, prevention and
PT treatment of a pathological condition resulting from a bacterial
PT infection.

XX
XX Example 1; SEQ ID NO 1399; 243bp; English.

XX The invention relates to an isolated nucleic acid derived from
CC Enterococcus faecium encoding an Enterococcus faecium polypeptide having
CC one of 10 fully defined sequences given in the (or comprising 40
CC sequential nucleotides chosen from any of the nucleic acids, its
CC complement or sequences hybridising to it). Also included are a
CC recombinant vector comprising the nucleic acid operably linked to
CC transcription regulatory element, a cell comprising the vector and a
CC single-stranded probe comprising the nucleic acid. The nucleic acids are
CC chosen from 3654 disclosed sequences encoding 3654 disclosed proteins.
CC The nucleic acids is useful for diagnosing pathological conditions
CC resulting from E. faecium bacterial infection (e.g. urinary tract
CC infection, bacteraemia, endocarditis, wounds and abdominal-pelvic
CC infection) and for screening drugs such as agonists and antagonists. The
CC nucleic acid is useful for recombinant production of Candida albicans -
CC derived peptides or antisense polypeptides. Pharmaceutical compositions
CC and vaccines containing the nucleic acid are useful for preventing or
CC treating Enterococcus faecium infections. The present sequence represents
CC one if the disclosed E. faecium nucleic acids.

SQ Sequence 1062 BP; 362 A; 204 C; 246 G; 250 T; 0 U; 0 Other;

Query March 2.9%; Score 139.8; DB 9; Length 1062;
Best Local Similarity 49.6%; Pred. No. 3,66-11;
Matches 484; Conservative 0; Mismatches 467; Indels 24; Gaps 4;

CY 645 TTTCAGTGCAGATCATTTGGGACAGCAATGACAAATTGCACTGTCAAGCATTGATAT 704
Db TTTTAGGACCAAGGCTCTTGCGGGTAAGCTTTGGGACAAAGTCTTCGCCGAATAACGACATG 115
CY 705 CAGTTAATCATATGGGGGACGTGATCATATGAATATATACATATATTAACACTTACCGAAAA 764
Db 116 ACCTTCGATTTGGGGGACATCGTGTCTTCAACAAGTAGAAGAATATGAACCCAACACGCA 175
CY 765 ATTAAAATATTTTACCACATATCATTTACCAAGCAACATATATGCAACGCAATATAG 824
Db 176 ATCAACCTTATTTTACCAGATCATTTATCTACCCGATTCATCCAAAGCAATTCCCAATATGG 235
CY 825 ACAGAGATATATCTGACACAAATATCATGTATTAATCTTAATATATCTTCCACACAAATATAC 884
Db 236 AAGATGCAATTAAGAAAGACGAGATGC---GGATGCTATTCGT-CATTCCAACGMAACAAATAC 292
CY 885 GCACCATATGTACCAAAATTCACACAAACAAACGACATATGTGTAAAAATATCTCAAATATTA 944
Db 293 GTTCTGTGTGCAAAAGAACTAGTGCCTAAA-----TTGAATATCAAAACGAGTGTATACC 346

QY 1209 TAATTCGATTCGATGGAATATATCGCGAAAAAATTAGTAATATGCTGTGCTA 1268
 Db 1260 TATTCACACAGTGTGAGAGAAATTTCTGAGGCGCTTGTTATGCGCATATGCTAAAGCGG 1319
 QY 1269 CTGTTAACTAAAGCAGATGAAATTAATAAACAATATATATAGCAAAAAATCATTCOA 1328
 Db 1320 CACTAATGACTCGCGGATGGAGAAATCATCTGCTCGGTGTC-----TGTGGCT 1373
 QY 1329 TACATCTTCAATATTAATTTGCTTCATCTGCTTGGAGATCTAATTAATACATGACA 1388
 Db 1374 CTATCCGCAAACTTTTACGATTTGACAGATTTGCTACTTAATCTGTTACTTGTACTA 1433
 QY 1389 CAGAACCTTACGCAATATGCTTTGACTAGAAATAGAAAGGTAGAAATATATAA 1448
 Db 1434 GTTGTGATTCACCTAATTTGGCGCTGCTGTAATATGCTAAGGCGGAAATTTAGATG 1493
 QY 1449 CATTAAATAGTCAACCTTAAGCTTTGTAAGGAACAGTACTGTAACCACTGATAT 1508
 Db 1494 AAGTATTTAGAAAAAATGGGCGATGCTGTGAAAGTGTTCGAAACAGCTAAAGCAGTACATG 1553
 QY 1509 CATTAGCAAAAAAATCTTAATGTAGAACTACCAATTTGCAATCTATCTAATTTATATAC 1568
 Db 1554 GTTGGCGGAAAAAATAGATATTTGATATGCGCAATTTACGAAATGATTTAGCGGATTTAT 1613
 QY 1569 ATGAGAAATATA 1579
 Db 1614 TCGAAAAATATA 1624

RESULT 39
 ABL3307/C
 ID ABL3307 standard; DNA; 6175 BP.

XX ABL33307;
 AC 26-MAR-2002 (first entry)
 DT
 XX Human immune system associated gene SEQ ID NO: 1280.
 DE
 XX Human; immune system disease; cytosine methylation; antiasthmatic;
 KM antiarteriosclerotic; antihaemic; cytosine; noctropic;
 KM neuroproliferative; anti-HIV; anticonvulsant; ophthalmological;
 KM antihemorrhagic; antidiabetic; antidiabetic; antipneumonic;
 KM antineuroinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
 KM acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 KM neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
 ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200200928-A2.
 PD
 XX 03-JAN-2002.
 PD
 XX 02-JUL-2001; 2001WO-EP007537;
 PF
 XX 30-JUN-2000; 2000DE-01032529.
 PR
 XX 01-SEP-2000; 2000DE-01043826.
 XX
 PA (EPIC-) EPICENOMICS AG.
 XX
 PI Olek A, Piepenbrock C, Berlin K;
 DR WPI; 2002-130909/17.
 XX
 PT Nucleic acid comprising fragment of chemically modified gene, useful for
 PT diagnosis and treatment of diseases associated with abnormal cytosine
 PT methylation.
 XX
 PS Claim 1; SEQ ID NO 1280; 32bp + Sequence Listing; German.
 CC The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences

CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 CC diseases. The present sequence is a gene of the invention
 XX
 SO Sequence 6175 BP; 1506 A; 154 C; 1357 G; 3158 T; 0 U; 0 Other;

Query Match 2.9%; Score 139; DB 6; Length 6175;
 Best Local Similarity 45.8%; Pred. No. 5.2e-13;
 Matches 800; Conservative 0; Mismatches 910; Indels 35; Gaps 8;

QY 740 ACACATATAACCTTACCGAAAAATTTAAATTTTACCCCATATGATCTACAGAC 799
 Db 6002 ATAAATATATAAACAATAATATATATTAACCTTTAAATATACAAATATATCTAAATAT 5943
 QY 800 AACATATATGCAACCAAGCAATATAGACAGATATTTATCTGACCAACATACATGATATATC 859
 Db 5942 TATATATAAACAACGATATTAACATACATAAATATATACATAAACAATAATATATATACAT 5883
 QY 860 TTAACCTTTCTTACACACATTTACGACCATTTGACCAATATACACACACACGAT 919
 Db 5882 ACGTAAACAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5823
 QY 920 ATGTGTAATAATTAATCAATATTTATTTGATGTAAGTAATGCAAAATTAATCACTCAAA 979
 Db 5822 TAAATATATAAACAACGATATTAACGATATTAATTAATTAATTAATTAATTAATTAAC 5767
 QY 980 TTTCCACGTAATATGACAGAGAAATTTTAAATATATATTAATTAATTAATTAATTAATTA 1039
 Db 5766 ATAAACATATATATTAACAAATATATATATATATATATATATATATATATATATATATA 5707
 QY 1040 CCAAGTTTCTTAAAGAAATTTGACAGACATCTCTGTTAGTATGATCTGCTGAT 1099
 Db 5706 TACAAATATATATTAACAAATATATATATATATATATATATATATATATATATATATA 5647
 QY 1100 AATTAAGAACTGTGATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1159
 Db 5646 TATTAATATATTAACATATATATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5587
 QY 1160 TACATCAAGATTTTATAGGTGACAGATTTGAGCGCATTAAGAAATTAATTTGCAAT 1219
 Db 5586 TACATATATTAATATTTAATCATATATATTAATTAATTAATTAATTAATTAATTAATTA 5527
 QY 1220 GCATGTGAATTAATCGCTGAAAAAATTTAGTAAATATGCTGCTGCTATTAAT 1279
 Db 5526 AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5467
 QY 1280 AAAGCATGAATGAATTTAAACATATATATATGCAAAAAATGATTCATGATCTTCAT 1339
 Db 5466 AAAGCATGAATGAATTTAAATATATATATATGCTAAATTAATTAATTAATTAATTAAT 5411
 QY 1340 ACAATTAATTTGCTCATGCTGCTGAGATCTAATTAATTAATTAATTAATTAATTAATTA 1399
 Db 5410 ACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5351
 QY 1400 CGCAATTTGCTTTGAGCATGAATTAAGAAAGGTAATTAATTAATTAATTAATTAATTA 1459
 Db 5350 AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5291
 QY 1460 CAGAACCTTAAGCTTTGTAAGAACAGATCTGTAACCACTGATATCATTAAGCAAA 1519
 Db 5290 ATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5231
 QY 1520 AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1579
 Db 5230 ATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5171
 QY 1580 TCACTAGATTAAGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1639
 Db 5170 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5112

QY 1640 CACATGATATAAAGGCTCAAAATATAGATACAGCT-TCACACATATATATAAA 1698
 Db 5111 AAAATATAACAAATATATATATATAAAGCTATATATAAATATATAAATATATAA 5052
 QY 1699 GCGCTCAATACCCATGTTACATCCAACTAGATAGAACATATTACATAAACAATT 1758
 Db 5051 TAAATATATAAAGCTATATATAAATATATATAAACAATATATAAATATATAA 4992
 QY 1759 AAGTATACCTAAATTTCCAGCTATATATATAAAGATTTCCATCTCCAT 1818
 Db 4991 AAAAATACATTAATATATAACCGTATATATAAACAATATATAATATATATAC 4932
 QY 1819 TAACCAATTGTTATCAGATATATATAATAGTCAAAATTTTACTTATTTATAAATA 1878
 Db 4931 TATATATATAAATATATATATAAATATATAATATATATATATAAATATATATAT 4875
 QY 1879 ATATGTTGACACTTTTTCATCATCATGATATATATATATATATATATATATATAT 1938
 Db 4874 AATATATATAAATATATATATATAAACAATATATATATATATATATATATATAA 4815
 QY 1939 AACTATATAAAGAACCAATACATATATATATAATTAACAACCTATTTACTACGTAAT 1998
 Db 4814 AATACATATATAATTTATAACATTAATATAAATATAAACAATATATAATATATATAT 4755
 QY 1999 ATTACATAGATGCTCAGCATTAATTAATTTATCATCTCTGACATATTTCAACATATA 2058
 Db 4754 AATACCATATATATAAACAATATATATAAATATATATAAATATATATATATATATAT 4700
 QY 2059 TATGATTAATCTGTTTATTTTATGCCCATTTAGTCTTATATAAATTTGCTAATTTACCCAAAA 2118
 Db 4699 ATAACTAAATACATTAATATAAATATATAAATATAAATATAAATATAAATATAA 4640
 QY 2119 GGTAAATACATCTCTGCAACATATAGAAAGATACATACAGTCAAAATATA 2178
 Db 4639 TATATAATATCAAAAT-----AAATATAATATAATATATATATATATATATATAT 4587
 QY 2179 AACAAATTTCTAATATTAACCACTAGCTTAAACATAGAACCAATAGTATATATATATAT 2238
 Db 4586 CACAAACATATATATAAACAATATATATAAATATATATATATATATATATATATAT 4537
 QY 2239 AGTATAGATATCCCTACAGCATGTATCAAAATATATAAAGAAAAATCGTATATATATAT 2298
 Db 4536 ATATATACATATATATATAAACAATATATATAAATATATATATATATATATATATAT 4477
 QY 2299 CAATTAACAAATTTCCAACTTCAACAACAACAATCTCAAGTCTATATATAAATTTCAAG 2358
 Db 4476 TAAATATATAAATATATAAATATATATATATATATATATATATATATATATATAT 4417
 QY 2359 AATATATATATATCTCAACAACAATATATATATATATATATATATATATATATATAT 2418
 Db 4416 AAAAATACGTTATATATAACGTAATATAAATATAAATATAAATATATATATATAT 4357
 QY 2419 AAAAAGTACCTCAGCATTTCAACATATATATATATATATATATATATATATATAT 2478
 Db 4356 AAAAAT 4297
 QY 2479 ATTAC 2483
 Db 4296 AATTAC 4292
 RESULT 40
 ABQ6998/c
 ID ABQ6998 standard; DNA; 37515 BP.
 XX
 AC ABQ6998;
 XX
 DT 28-AUG-2002 (first entry)
 XX
 DE Human angiogenesis associated polynucleotide SEQ ID NO 28.
 XX
 KW Human; angiogenesis; methylation; eye disease; glaucoma; tumour;
 inflammation; rheumatoid arthritis; diabetic retinopathy; anti-infects;

KW macular degeneration; inflammatory bowel disease; Crohn's disease;
 KW anti-rheumatic; anti-rheumatic; anti-diabetic; anti-psoriasis;
 KW anti-rheumatic; anti-rheumatic; anti-diabetic; anti-psoriasis;
 XX
 OS Homo sapiens.
 XX
 PN WO200246454-A2.
 XX
 PD 13-JUN-2002.
 XX
 PF 06-DEC-2001; 2001WO-EP014320.
 XX
 PR 06-DEC-2000; 2000DE-01061338.
 XX
 PA (EPIC-) EPIGENOMICS AG.
 XX
 PI Schacht O;
 XX
 DR WPI; 2002-500450/53.
 XX
 PT New nucleic acid fragments from chemically treated angiogenesis-
 PT associated genes, useful for determining methylation status, e.g. in
 PT diagnosis or treatment of cancer.
 XX
 PS Claim 1; SEQ ID NO 28; 41bp + Sequence Listing; German.
 XX
 CC The invention relates to a nucleic acid (I) comprising a segment of 18
 CC bases of chemically pretreated DNA of angiogenesis-associated genes (II)
 CC having sequences (ABQ6997-ABQ6998) or their complements. (II) also
 CC related oligomers, are used to evaluate the methylation status and/or
 CC single-nucleotide polymorphisms, in angiogenesis-related genes, for
 CC diagnosis and treatment of eye diseases, proliferative retinopathy,
 CC neovascular glaucoma, solid tumours, inflammation, rheumatoid arthritis,
 CC diabetic retinopathy, macular degeneration caused by neovascularisation,
 CC psoriasis, arteriosclerosis, inflammatory bowel diseases, ulcers and
 CC Crohn's disease. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pat_sequences
 XX
 SQ Sequence 37515 BP; 11428 A; 425 C; 7165 G; 18497 T; 0 U; 0 Other;
 Query Match 2.9%; Score 138.8; DB 6; Length 37515;
 Best Local Similarity 42.5%; Pred. No. 68-13;
 Matches 1853; Conservative 0; Mismatches 2437; Indels 72; Gaps 18;
 QY 165 AGAAGCAACCAATTTCTCAATCTCCAAATTTTACATGCAAGATTTATATATATAT 224
 Db 18848 AAACACATACGATATTTTAAACATATACAAAAAACCACATTCATCTCTCTCATATA 18789
 QY 225 AGAAGCAACCAATTTCTCAATCTCCAAATTTTACATGCAAGATTTATATATATAT 284
 Db 18788 AATACATACGATCTCTCAATCTCTCAATCTCTCAATCTCTCAATCTCTCAATCTCT 18730
 QY 285 AACCACTGATTCGAAGATACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 344
 Db 18729 AAAAACCACCAAAAAATCTACTCTCAATTAACCTTTCTCTCTCTCTCTCTCTCTCT 18670
 QY 345 AATATCATAGAGTCTTTATTTGCAATACATATATAAATCAATCAAAACCAAAATATATAT 404
 Db 18669 CATTTAACTA-ATTATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 18611
 QY 405 ACTATCTAATTAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 464
 Db 18610 ACCCTCTATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 18551
 QY 465 AACTTACTATTAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 524
 Db 18550 AATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 18491
 QY 525 TCACATCTGACATGAT-TAAACAATAGTACACATTTCTTAACTAGATATATATATAT 583
 Db 18490 CTTCAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 18431

[illegible][illegible]

Search completed: April 5, 2004, 13:26:52
Job time : 1773 secs

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2004, 10:55:36 ; Search time 17898 Seconds

(without alignments) 11502.928 Million cell updates/sec

Title: US-10-081-051-52

Perfect score: 4750

Sequence: 1 gatccctgctactactttagg.....ccacatgactaatgac 4750

Scoring table: IDENTITY_NUC

Gapop 10.0 ; Gapext 1.0

Searched: 3470272 seqs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenBank1.*

1: gb_da.*

2: gb_hvg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_da.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_ov.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_hvg_hum.*

31: em_hvg_inv.*

32: em_hvg_other.*

33: em_hvg_mus.*

34: em_hvg_pln.*

35: em_hvg_rdd.*

36: em_hvg_man.*

37: em_hvg_vtc.*

38: em_sy.*

39: em_hvg_hum.*

40: em_hvg_mus.*

41: em_hvg_other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4750	100.0	4750	1	AF308670
2	4750	100.0	4750	1	AF308670
3	1416	29.8	1416	6	AX685249
4	984	20.7	984	6	AX685247
5	597	12.6	597	6	AX685250
6	371	7.8	371	1	AX685246
7	357.6	7.5	279110	1	REPX03
8	338.6	7.1	12477	1	AE008640
9	231	4.9	23906	1	AF176314
10	219.6	4.6	14211	1	AE011408
11	216	4.5	11400	1	AB080665
12	207.8	4.4	34980	6	AX344555
13	204	4.0	349174	1	AB063522
14	188.6	4.0	4362	1	PVGLNABC
15	186.4	3.9	8056	6	AX599046
16	184	3.9	4601	3	DMU11584
17	184	3.9	19517	3	DMU37541
18	183	3.9	172816	9	AC093899
19	182.2	3.8	14867	3	AE001398
20	182	3.8	300205	1	AE015944
21	181.4	3.8	145992	9	AE015939
22	180.2	3.8	122168	9	AE015938
23	175.8	3.7	12277	1	AE010598
24	174.4	3.7	333750	1	AP004827
25	173.8	3.7	104992	2	AC005504
26	173.8	3.7	169546	2	AC004157
27	173.8	3.7	250421	2	AE014849
28	173.4	3.7	301278	1	AE015939
29	172.8	3.6	301550	1	AE003134
30	172.8	3.6	303414	1	AE015938
31	172.8	3.6	346900	1	AP003362
32	170.8	3.6	282183	1	CJ11168X4
33	170.6	3.6	1425	6	AR377574
34	169.2	3.6	10861	1	AE007680
35	169	3.6	3900	1	AY309266
36	168.8	3.6	13877	2	AC120883
37	168.8	3.6	349751	3	PFMAL4P3
38	168.6	3.5	996	6	AX622666
39	168.4	3.5	67970	3	PFMAL1P3
40	167	3.5	19517	3	DMU37541
41	166.8	3.5	17559	2	AC145101
42	166.4	3.5	13124	1	AE008105
43	166.4	3.5	13147	1	AE008070
44	166	3.5	254050	3	PFA929358
45	165.8	3.5	10277	1	AE001176

ALIGNMENTS

RESULT 1

LOCUS AF308670

DEFINITION AF308670 4750 bp DNA linear BCT 03-OCT-2001

gene, partial cds; hypothetical glycerol 3-phosphate dehydrogenase and hypothetical nitrogen assimilation regulatory protein genes, complete cds; and hypothetical cell division protein fteq gene, partial cds.

ACCESSION AF308670

VERSION AF308670.1 GI:15811157

SOURCE

ORGANISM

Enrichia ruminantium (heartwater rickettsia)

Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales; Anaplasmataceae; Enrichia.

Pred. No. is the number of results predicted by chance to have a

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCES
AUTHORS
TITLE
JOURNAL
FEATURES
SOURCE

1 (bases 1 to 4750)
Barbet,A.F., Whitmore,M.M., Kamper,S.M., Simbi,B.H., Ganta,R.R.,
Moreland,A.L., Mwambi,D.M., McGuire,T.C. and Mahan,S.M.
A subset of Cowdria ruminantium genes important for immune
recognition and protection
Gene 275 (2), 287-298 (2001)
21472264
11587856

2 (bases 1 to 4750)
Barbet,A.F., Whitmore,M.M., Kamper,S.M., Simbi,B.H., Ganta,R.R.,
Moreland,A.L., Mwambi,D.M., McGuire,T.C. and Mahan,S.M.
Direct Submission
Submitted (26-SEP-2000) Department of Pathobiology, University of
Florida, P.O. Box 110980, Gainesville, FL 32611-0880, USA
Location/Qualifiers
1. 4750
/organism="Ehrlichia ruminantium"
/mol_type="genomic DNA"
/strain="Highway"
/db_xref="taxon:779"
/clone="19hw"
complement (<1. 371)
/note="orf11"
/codon_start=1
/evidence=not experimental
/transl_table=1
/product="hypothetical integrase/recombinase"
/protein_id="AA08833.1"
/db_xref="GI:15811158"
/translation="MSNKLVDYIVNNWKIMWEGEKYSINTVSVYRDLKTEFLHK
STGCVTLNVLVHKIGDLKRWPSRQCVQVEAVTNARSLSALRNFRYSIMYYIDS
QAVFYLSRPHLKKTLPLVLAGS"
632..1615
/note="orf2"
/codon_start=1
/evidence=not experimental
/transl_table=1
/product="hypothetical glycerol 3-phosphate dehydrogenase"
/protein_id="AA08834.1"
/db_xref="GI:15811159"
/translation="MKISILGAGSPGATAIALSAGISVNLGRDHRNITHNYRK
NLYLPPYHLPDNIVATSNIDEVSDNNTCIITPTCQLEPTCTQCHKQMKQNP
ILISKEIETSLKPSREAEILQYRPIISGSPSAKEIAEHLFPCSVLQDNKEE
GESILFETISNDVKIITYHODIGVQIGALAKNIIAICGIIAGKMLGNNAVAIVITG
MNEIKTLIYAKNSIDLTLLIGSCGLDILCTCTESRNMAFGLEIGKGRINILID
HNKLVEGTSTVPLSLAKLNVELPICIISYVNLHENISDKAISNLS"
complement (2467..3882)
/note="orf3"
/codon_start=1
/evidence=not experimental
/transl_table=1
/product="hypothetical nitrogen assimilation regulatory
protein"
/protein_id="AA08835.1"
/db_xref="GI:15811160"
/translation="MAQDFEMSKRLYISVLVVDDEVDIRNLKILISDNYVTKLA
VDDLSATKMAVEKEPDDVLLDIWLRSDIDGLSVLEKERYPYLPVIMTSKGLAT
AVKSLHWGADVIEKPTSEGRKLAVYKRAIESGRLEKRENDKSAFEDIEIYGNFVI
RNLRSMNKKAATISSRLITIGSGVEKVAARLHKSKGYDPTFSWSSMLPANNY
LVNI FGSSENNILSHRPVPHIGIEQANHGTLFIDEVDLRYDLRLRLLOGKI
LVNSKIPVSDIVRIIVSSKDISEVKAQGFCEDLVYRLNLPTRVPELVAYCTDIP
ELCRVYNSICRKGICHTVLSDEALIAQSYEMPNLQNLNVEIMILIMSPREMI
TAKDLPDVISNPINDVLSAKYISVPLKARREPEROYLKQLSRFGNSVRTAFV
GMSRSLHRLKIGLCNVSE"
complement (4154..4750)
/note="orf4"
/codon_start=1
/evidence=not experimental
/transl_table=1
/product="hypothetical cell division protein fteq"
/protein_id="AA08836.1"
/db_xref="GI:15811161"
/translation="DQLVNCGFTVDKILIDGNEYSSDEIRKLVDARSIFVPLADLR

ORIGIN

Query Match 100.0%; Score 4750; DB 1; Length 4750;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4750; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GATCGTGTAACTAGTAAATGTTTTTTTAAAGATGAGCGCTAGATAAATAAAGACT 60
1 GATCGTGTAACTAGTAAATGTTTTTTTAAAGATGAGCGCTAGATAAATAAAGACT 60
61 GCTGACTGTCTATATAGTACATACGTGATGATGATGATGATGATGATGATGATG 120
61 GCTGACTGTCTATATAGTACATACGTGATGATGATGATGATGATGATGATGATG 120
121 AGTGACCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180
121 AGTGACCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180
181 CTCAATCTCCAAATTTTAAACATGACAAAGATTATTATGTAACAGAACACCTGTA 240
181 CTCAATCTCCAAATTTTAAACATGACAAAGATTATTATGTAACAGAACACCTGTA 240
241 TTATGCAAGAACTGATTAATCTTCAAGGCTCTTCAATGATGATGATGATGATGAT 300
241 TTATGCAAGAACTGATTAATCTTCAAGGCTCTTCAATGATGATGATGATGATGAT 300
301 GAATACCTCTCTCTGTTCCATCATATTTTCCAAATTTTAAATATCATAGAGTTCT 360
301 GAATACCTCTCTCTGTTCCATCATATTTTCCAAATTTTAAATATCATAGAGTTCT 360
361 TTATGCAAGAACTGATTAATCTTCAAGGCTCTTCAATGATGATGATGATGATGAT 420
361 TTATGCAAGAACTGATTAATCTTCAAGGCTCTTCAATGATGATGATGATGATGAT 420
421 AAAATCTCTCAATAGTCTTTGACAAAGATCAATATTAATGACACCTTAATATTA 480
421 AAAATCTCTCAATAGTCTTTGACAAAGATCAATATTAATGACACCTTAATATTA 480
481 TATATTTCAATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 540
481 TATATTTCAATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 540
541 TAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
541 TAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
601 GACTGCTTTCTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 660
601 GACTGCTTTCTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 660
661 ATTGGGACAGCAATGACATGACATGACATGACATGACATGACATGACATGACATG 720
661 ATTGGGACAGCAATGACATGACATGACATGACATGACATGACATGACATGACATG 720
721 ACGTGATCATGAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 780
721 ACGTGATCATGAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 780
781 CACATATCATATGACAAACATATGACAAACATATGACAAACATATGACAAACATAT 840
781 CACATATCATATGACAAACATATGACAAACATATGACAAACATATGACAAACATAT 840
841 CAACATATCATATGACAAACATATGACAAACATATGACAAACATATGACAAACATAT 900
841 CAACATATCATATGACAAACATATGACAAACATATGACAAACATATGACAAACATAT 900
901 AATCAACACAAACAGATATGATGATGATGATGATGATGATGATGATGATGATGAT 960
901 AATCAACACAAACAGATATGATGATGATGATGATGATGATGATGATGATGATGAT 960

QY 961 CGAATTCATCAGTCAATTTCCAGTGAATAGCAGAAATTTTACATATATCC 1020
Db 961 CGAATTCATCAGTCAATTTCCAGTGAATAGCAGAAATTTTACATATATCC 1020
QY 1021 AATTTTACTCTCTGCTGCAAGTTTTCAGAAATTCAGAAACATCTCTGCTGAG 1080
Db 1021 AATTTTACTCTCTGCTGCAAGTTTTCAGAAATTCAGAAACATCTCTGCTGAG 1080
QY 1081 TTTACTACTGCTGCTGCAATTAAGAACTGCTGCAATTCATGATGAATTAAGTAA 1140
Db 1081 TTTACTACTGCTGCTGCAATTAAGAACTGCTGCAATTCATGATGAATTAAGTAA 1140
QY 1141 TGATGTTCTAAATTAATATACATCAAGATTTATAGTGTACAGATTTGAGCTGAT 1200
Db 1141 TGATGTTCTAAATTAATATACATCAAGATTTATAGTGTACAGATTTGAGCTGAT 1200
QY 1201 AAGAAACATTAATTCGAATTCGATGGAATTAATGCTGGAATTAATTAATTC 1260
Db 1201 AAGAAACATTAATTCGAATTCGATGGAATTAATGCTGGAATTAATTAATTC 1260
QY 1261 TGTGCTACTGTTATTAACCTAAAGCATGATGAAATTAACATATATAGCAAAA 1320
Db 1261 TGTGCTACTGTTATTAACCTAAAGCATGATGAAATTAACATATATAGCAAAA 1320
QY 1321 TCATTCATATAGATCTTCAATCAATTAATGCTGATCATGCTGCTGAGATCTAAATTTAC 1380
Db 1321 TCATTCATATAGATCTTCAATCAATTAATGCTGATCATGCTGCTGAGATCTAAATTTAC 1380
QY 1381 ATGTCACACAGACATTCACGCAATATGCTTTTGAATAGAAATAGGAAAGGTAGAAA 1440
Db 1381 ATGTCACACAGACATTCACGCAATATGCTTTTGAATAGAAATAGGAAAGGTAGAAA 1440
QY 1441 TATATAATCATTAATATAGATCAACACCTTAAGCTTGTGAGAAACAGATCTGTAACC 1500
Db 1441 TATATAATCATTAATATAGATCAACACCTTAAGCTTGTGAGAAACAGATCTGTAACC 1500
QY 1501 ACTGATATCATAGGAAATTAATATAGTGAATCAACCTTTGCTATATCTATTTTCA 1560
Db 1501 ACTGATATCATAGGAAATTAATATAGTGAATCAACCTTTGCTATATCTATTTTCA 1560
QY 1561 TTTATATCATGAGATATATCATGATTAAGCAATCAACATTAATTTAGTCTAT 1620
Db 1561 TTTATATCATGAGATATATCATGATTAAGCAATCAACATTAATTTAGTCTAT 1620
QY 1621 CATCTTTGATATTTTCAACACATGATTAATTAACGCTCAAAATATAGATCAACGT 1680
Db 1621 CATCTTTGATATTTTCAACACATGATTAATTAACGCTCAAAATATAGATCAACGT 1680
QY 1681 TCACACATATATAAAGAGAGCTCAAAATCCATGCTTACATCACTAAATTAATTAAC 1740
Db 1681 TCACACATATATAAAGAGAGCTCAAAATCCATGCTTACATCACTAAATTAATTAAC 1740
QY 1741 ATATTAATCAATCAAACTTAAGTATTAATTAATTTCTAGCTAATTAATTAACAAAT 1800
Db 1741 ATATTAATCAATCAAACTTAAGTATTAATTAATTTCTAGCTAATTAATTAACAAAT 1800
QY 1801 CCTATCTCTATCTCAATTAACCAATTTTATGATTAATTAATTAATTAATTAACCT 1860
Db 1801 CCTATCTCTATCTCAATTAACCAATTTTATGATTAATTAATTAATTAATTAACCT 1860
QY 1861 TTTACTATTTTAAATTAATTAATTTTGAACCTTTTCACTACATCAATTAATTAATTAAC 1920
Db 1861 TTTACTATTTTAAATTAATTAATTTTGAACCTTTTCACTACATCAATTAATTAATTAAC 1920
QY 1921 CATATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAC 1980
Db 1921 CATATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAC 1980
QY 1981 ACCTATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAC 2040
Db 1981 ACCTATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAC 2040
QY 2041 CATATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAC 2100
Db 2041 CATATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAC 2100
QY 2101 TTTGCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAC 2160
Db 2101 TTTGCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAC 2160
QY 2161 TACAGCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAC 2220
Db 2161 TACAGCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAC 2220
QY 2221 AATAGATATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAC 2280
Db 2221 AATAGATATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAC 2280
QY 2281 TGTGTTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAC 2340
Db 2281 TGTGTTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAC 2340
QY 2341 TATATTAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAC 2400
Db 2341 TATATTAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAC 2400
QY 2401 ATACATCATATCTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAC 2460
Db 2401 ATACATCATATCTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAC 2460
QY 2461 CAATGTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAC 2520
Db 2461 CAATGTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAC 2520
QY 2521 AGCTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAC 2580
Db 2521 AGCTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAC 2580
QY 2581 AGTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAC 2640
Db 2581 AGTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAC 2640
QY 2641 TTTAGCACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAC 2700
Db 2641 TTTAGCACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAC 2700
QY 2701 TGTGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAC 2760
Db 2701 TGTGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAC 2760
QY 2761 TGTGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAC 2820
Db 2761 TGTGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAC 2820
QY 2821 TACATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAC 2880
Db 2821 TACATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAC 2880
QY 2881 CGGTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAC 2940
Db 2881 CGGTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAC 2940
QY 2941 ATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAC 3000
Db 2941 ATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAC 3000
QY 3001 AGACACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAC 3060
Db 3001 AGACACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAC 3060
QY 3061 TGTGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAC 3120
Db 3061 TGTGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAC 3120
QY 3121 ATCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAC 3180
Db 3121 ATCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAC 3180

Db 3121 ATGTAATTAATAGTACATGATTTGCTGCTCTAATATTCGAATATGAGAGTACTCT 3180
 QY 3181 ATGAGACATAATATTTACTTTCTCTACACCAAAATATTTACCAAGTAAATTTAGC 3240
 Db 3181 ATGAGACATAATATTTACTTTCTCTACACCAAAATATTTACCAAGTAAATTTAGC 3240
 QY 3241 TGGAGACATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAG 3300
 Db 3241 TGGAGACATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAG 3300
 QY 3301 TATTAAGCTTACCACT 3360
 Db 3301 TATTAAGCTTACCACT 3360
 QY 3361 CGATGTAGTACGCT 3420
 Db 3361 CGATGTAGTACGCT 3420
 QY 3421 GACTATTTCAATATCT 3480
 Db 3421 GACTATTTCAATATCT 3480
 QY 3481 AGACTATAGCT 3540
 Db 3481 AGACTATAGCT 3540
 QY 3541 ATATCATTAAGCAAGCAATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3600
 Db 3541 ATATCATTAAGCAAGCAATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3600
 QY 3601 AATCATTAATAGCAAGCAATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3660
 Db 3601 AATCATTAATAGCAAGCAATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3660
 QY 3661 ATCAATATAGCAATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3720
 Db 3661 ATCAATATAGCAATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3720
 QY 3721 AGCATATGATAGCAATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3780
 Db 3721 AGCATATGATAGCAATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3780
 QY 3781 TAAATATCTTTATGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGAT 3840
 Db 3781 TAAATATCTTTATGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGAT 3840
 QY 3841 AATATACATCTTTCT 3900
 Db 3841 AATATACATCTTTCT 3900
 QY 3901 TGAATATCTGCAATATGAG 3960
 Db 3901 TGAATATCTGCAATATGAG 3960
 QY 3961 CTATGATCAACCAATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4020
 Db 3961 CTATGATCAACCAATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4020
 QY 4021 TATTAATATGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGAT 4080
 Db 4021 TATTAATATGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGAT 4080
 QY 4081 AAGAAATATGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGAT 4140
 Db 4081 AAGAAATATGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGAT 4140
 QY 4141 AATTAATATGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGAT 4200
 Db 4141 AATTAATATGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGAT 4200
 QY 4201 TCCATATGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAG 4260
 Db 4201 TCCATATGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAG 4260

QY 4261 AAGAATCATTTATGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAG 4320
 Db 4261 AAGAATCATTTATGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAG 4320
 QY 4321 GACTATCAACATATGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAG 4380
 Db 4321 GACTATCAACATATGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAG 4380
 QY 4381 CTTCAGATATTAATCTGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGAT 4440
 Db 4381 CTTCAGATATTAATCTGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGAT 4440
 QY 4441 CCCGATATATCAATATGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAG 4500
 Db 4441 CCCGATATATCAATATGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAG 4500
 QY 4501 TATCATGATACCAATTTGCAAAAGCAAGATATTTCTGATAGATAGATAGATAGATAG 4560
 Db 4501 TATCATGATACCAATTTGCAAAAGCAAGATATTTCTGATAGATAGATAGATAGATAG 4560
 QY 4561 GTAACAGCTCTTAACTAG 4620
 Db 4561 GTAACAGCTCTTAACTAG 4620
 QY 4621 TCAATCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 4680
 Db 4621 TCAATCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 4680
 QY 4681 CAGAACTTACATATGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAG 4740
 Db 4681 CAGAACTTACATATGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAG 4740
 QY 4741 CTAATGATC 4750
 Db 4741 CTAATGATC 4750

RESULT 2
 A685245
 LOCUS A685245 4750 bp DNA linear PAT 29-MAR-2003
 DEFINITION Sequence 52 from Patent WO0206652.
 A685245
 ACCESSION A685245
 VERSION A685245.1 GI:29371587
 KEYWORDS
 SOURCE Ehrlichia ruminantium (heartwater rickettsia)
 ORGANISM Ehrlichia ruminantium
 Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
 Anaplasmataceae; Ehrlichia.
 REFERENCE
 1 Barber, A.F., Whitlure, W.M., Kamper, S.M., Simbi, B.H., Ganta, R.R.,
 Moreland, A.L., Mwambi, D.M., McGuire, T.C. and Mahan, S.M.
 Ehrlichia ruminantium polypeptides, antigens, polynucleotides, and
 methods of use
 Patent: WO 0206652-A 52 29-AUG-2002;
 JOURNAL University of Florida (US)
 FEATURES
 source
 1. 4750
 /organism="Ehrlichia ruminantium"
 /mol_type="unassigned DNA"
 /db_xref="taxon:779"
 ORIGIN
 Query Match 100.0%; Score 4750; DB 6; Length 4750;
 Best local Similarity 100.0%; P-adj. No. 0;
 Matches 4750; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GATCTGTATATGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGAT 60
 Db 1 GATCTGTATATGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGAT 60
 QY 61 GCTTACTGTCTATATGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAG 120
 Db 61 GCTTACTGTCTATATGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAG 120

Db 61 GCTGACTGCTCTATATAGTACATACGTGATGATAGCAAAAAATTTCTCAAGACAGAC 120
Qy 121 AGTACCTTGATGATGTAAGTCTTCACTCACTCATGATATCTAGAGCAAACTTTT 180
Db 121 AGTACCTTGATGATGTAAGTCTTCACTCACTCATGATATCTAGAGCAAACTTTT 180
Qy 181 CTCAATCTGCAATTTTACATGACAGATATTTATGTAACAGAACCTCTGACTC 240
Db 181 CTCAATCTGCAATTTTACATGACAGATATTTATGTAACAGAACCTCTGACTC 240
Qy 241 TTATGACAGAACTCTATTAACCTTCAAGGCTCTTACATAGAAACATGATTCAA 300
Db 241 TTATGACAGAACTCTATTAACCTTCAAGGCTCTTACATAGAAACATGATTCAA 300
Qy 301 GAATACCTCTCTGCTGTCATCAATTTTCCATATTTTCAATATCATAGAGTCT 360
Db 301 GAATACCTCTCTGCTGTCATCAATTTTCCATATTTTCAATATCATAGAGTCT 360
Qy 361 TTATTCGACATACATTTTAAATCAATCAAAACAAATATATGATCATATTAATA 420
Db 361 TTATTCGACATACATTTTAAATCAATCAAAACAAATATATGATCATATTAATA 420
Qy 421 AAAAATCCTCAATAGTCTTTAGCAGAACTACAAATATTAATGACACTTATTTAA 480
Db 421 AAAAATCCTCAATAGTCTTTAGCAGAACTACAAATATTAATGACACTTATTTAA 480
Qy 481 TATATTCGAATTTTATTTATCTACTCATATTTTACATATTTACATCTGACATGAT 540
Db 481 TATATTCGAATTTTATTTATCTACTCATATTTTACATATTTACATCTGACATGAT 540
Qy 541 TAAACATAGTAAACATTTCTTAAACATAGATATATATCAATATTAATTAATTTT 600
Db 541 TAAACATAGTAAACATTTCTTAAACATAGATATATATCAATATTAATTAATTTT 600
Qy 601 GACTGCTTTCTGATATTAATTTAGTAAATGAAATAGATTTTAGTGCAGAGATC 660
Db 601 GACTGCTTTCTGATATTAATTTAGTAAATGAAATAGATTTTAGTGCAGAGATC 660
Qy 661 ATTGCGACAGCAATAGCAATGCTGCTGACGACATGATATCACTTAACTTATGAG 720
Db 661 ATTGCGACAGCAATAGCAATGCTGCTGACGACATGATATCACTTAACTTATGAG 720
Qy 721 ACGTATCATAGAAATTTACATATTAACACTTACCGAAATTTTAAATTTTACC 780
Db 721 ACGTATCATAGAAATTTACATATTAACACTTACCGAAATTTTAAATTTTACC 780
Qy 781 CACATCATCTCTACAGCAACATATATGCAACGAGATATAGAGGATATTAATCTGA 840
Db 781 CACATCATCTCTACAGCAACATATATGCAACGAGATATAGAGGATATTAATCTGA 840
Qy 841 CAACATATCATATTTATTTCTTAACTATTTCTACACAACTTACCGACATATGACACA 900
Db 841 CAACATATCATATTTATTTCTTAACTATTTCTTAACTATTTCTTAACTATTTAC 900
Qy 901 AATACACAGAAACAGATATGTAATTAATCTCCAAATTTTAAATTTTAAAGTAT 960
Db 901 AATACACAGAAACAGATATGTAATTAATCTCCAAATTTTAAATTTTAAAGTAT 960
Qy 961 CGAATATCATCTCAATTTTCCAGTGAATATGAGAGAAATTTTCAATATATATCC 1020
Db 961 CGAATATCATCTCAATTTTCCAGTGAATATGAGAGAAATTTTCAATATATATCC 1020
Qy 1021 AATTTTATATCTCTGCTGCTCAAGTTTGTCTAAAGAAATTTGAGAACTCTCTGTA 1080
Db 1021 AATTTTATATCTCTGCTGCTCAAGTTTGTCTAAAGAAATTTGAGAACTCTCTGTA 1080
Qy 1081 TATAGTCTTCTGCTGCTGATATTAAGAACTTGTGATCATGATGAAACATATAGTAA 1140
Db 1081 TATAGTCTTCTGCTGCTGATATTAAGAACTTGTGATCATGATGAAACATATAGTAA 1140
Qy 1141 TGAATTTCTTAAATATATATCAATCAAGATATTTATGCTGTAACAGATGAGCTGAT 1200
Db 1141 TGAATTTCTTAAATATATATCAATCAAGATATTTATGCTGTAACAGATGAGCTGAT 1200

Qy 1201 AAAGAACATTAATGCAATTTGATGAGATATATCGCTGAGAAAAATTTAGGTAATATGC 1260
Db 1201 AAAGAACATTAATGCAATTTGATGAGATATATCGCTGAGAAAAATTTAGGTAATATGC 1260
Qy 1261 TGTGCTCTGCTGATATTAATTAAGGATGATGAAATTTAAACCTATATATAGCAAAAA 1320
Db 1261 TGTGCTCTGCTGATATTAATTAAGGATGATGAAATTTAAACCTATATATAGCAAAAA 1320
Qy 1321 TCATTCATATAGATCTTCAATTAATTTGCTCATCATGCTCTGAGAGCTATATTTAC 1380
Db 1321 TCATTCATATAGATCTTCAATTAATTTGCTCATCATGCTCTGAGAGCTATATTTAC 1380
Qy 1381 ATGACAAAGAACATTCACGCAATATGCTTTGAGACTAGAAAATGAGAAAGGTAGAA 1440
Db 1381 ATGACAAAGAACATTCACGCAATATGCTTTGAGACTAGAAAATGAGAAAGGTAGAA 1440
Qy 1441 TATAATTCATTAATATAGATCAACGCTTAAAGCTTTGAGAGAAACGATATGTAACC 1500
Db 1441 TATAATTCATTAATATAGATCAACGCTTAAAGCTTTGAGAGAAACGATATGTAACC 1500
Qy 1501 ACTGATATCATATAGCAAAAAAATTTATGTAAGAACTACAAATTTGATATTTTAA 1560
Db 1501 ACTGATATCATATAGCAAAAAAATTTATGTAAGAACTACAAATTTGATATTTTAA 1560
Qy 1561 TTTATTAATAGAGAAATATATGATAGTAAGAAAGCAATATTAATTTATGCTGAT 1620
Db 1561 TTTATTAATAGAGAAATATATGATAGTAAGAAAGCAATATTAATTTATGCTGAT 1620
Qy 1621 CATCTTTGATATTTTCAACCAATGATTAATTAACGCTCAAAAAATATAGAAATCAAC 1680
Db 1621 CATCTTTGATATTTTCAACCAATGATTAATTAACGCTCAAAAAATATAGAAATCAAC 1680
Qy 1681 TCACGACATATTAATTAAGAGCTCAATATCCCATGCTTAACATCACTAATGATGTAAC 1740
Db 1681 TCACGACATATTAATTAAGAGCTCAATATCCCATGCTTAACATCACTAATGATGTAAC 1740
Qy 1741 ATATACATTAACCAACTTAATTAATTAATTTCTGATATTAATTAATTAACAAAT 1800
Db 1741 ATATACATTAACCAACTTAATTAATTAATTTCTGATATTAATTAATTAACAAAT 1800
Qy 1801 CCTATCTCTATCTCATATTAACCAATTTGTAATGATATTAATTAATTAACAAATTT 1860
Db 1801 CCTATCTCTATCTCATATTAACCAATTTGTAATGATATTAATTAATTAACAAATTT 1860
Qy 1861 TTAATTTTATTAATTAATTAATTTGCACTTTTCACTACATCATGATATTAATGAA 1920
Db 1861 TTAATTTTATTAATTAATTAATTTGCACTTTTCACTACATCATGATATTAATGAA 1920
Qy 1921 CATATTTTCAATTAATTAATTAATTAATTAAGAACCAATATCATATTAATTAATTAAC 1980
Db 1921 CATATTTTCAATTAATTAATTAATTAATTAATTAAGAACCAATATCATATTAATTAAC 1980
Qy 1981 ACCATATTAACAGAAATATTAACATAGTCTGCAATTAATTAATTTTCACTCTGTA 2040
Db 1981 ACCATATTAACAGAAATATTAACATAGTCTGCAATTAATTAATTTTCACTCTGTA 2040
Qy 2041 CATATTTTCAACATTAATTAATTAATTAATTTTATGCTTATGCTTATTAATTA 2100
Db 2041 CATATTTTCAACATTAATTAATTAATTAATTTTATGCTTATGCTTATTAATTA 2100
Qy 2101 TTGCTTAATTAACCAAAAAAGTTTAAATTAACCATCTGCAACATTAATTAAGAAATCA 2160
Db 2101 TTGCTTAATTAACCAAAAAAGTTTAAATTAACCATCTGCAACATTAATTAAGAAATCA 2160
Qy 2161 TACAGCTCAATTAATTAATTAATTAATTTCTATATTAACCACTTCTTAAACATGGAAC 2220
Db 2161 TACAGCTCAATTAATTAATTAATTAATTTCTATATTAACCACTTCTTAAACATGGAAC 2220
Qy 2221 AATAGTATATCATATCTTAAGTATATCTTACAGATGATCAATTAATTAAGAAATTA 2280
Db 2221 AATAGTATATCATATCTTAAGTATATCTTACAGATGATCAATTAATTAAGAAATTA 2280

QY	2281	TCGTTATGTAATTAATTAATTCACATTTCAAAATTTCCAACTCCAAACAACAATCTCAAAAGTTC	23408
Db	2281	TCGTTATGTAATTAATTAATTCACATTTCAAAATTTCCAACTCCAAACAACAATCTCAAAAGTTC	23408
QY	2341	TATTAATAACAATTCAGAAATTTATATATATACCTTCACCAACAATAGTATTCAAATTCCTTA	24000
Db	2341	TATTAATAACAATTCAGAAATTTATATATATACCTTCACCAACAATAGTATTCAAATTCCTTA	24000
QY	2401	ATACATCATATCCCTTTTAAAAAGATGACCTTCACAAATTCATACATCAATTAATTAATACATG	24606
Db	2401	ATACATCATATCCCTTTTAAAAAGATGACCTTCACAAATTCATACATCAATTAATTAATACATG	24606
QY	2461	CAATGTTATTCAGAAACATTTACACAAATTCAGAAATTTCAATTTACGCTGTAATGCTGA	25202
Db	2461	CAATGTTATTCAGAAACATTTACACAAATTCAGAAATTTCAATTTACGCTGTAATGCTGA	25202
QY	2521	ACGTTCCATCCCAACAATTCAGCAAGTTCGATATCATTTACCTCCAAACAGAGATTAACG	25808
Db	2521	ACGTTCCATCCCAACAATTCAGCAAGTTCGATATCATTTACCTCCAAACAGAGATTAACG	25808
QY	2581	AGTTTAAAAATACGTCTTTCAAAATTCCTCAGACCTTACGTAATGCTACAGAAATTAAC	26404
Db	2581	AGTTTAAAAATACGTCTTTCAAAATTCCTCAGACCTTACGTAATGCTACAGAAATTAAC	26404
QY	2641	TTTAGCACTTAAACATCATTAATTAGCGAAATTAGATCTATATCTATCTATCTGTAATCTTT	27000
Db	2641	TTTAGCACTTAAACATCATTAATTAGCGAAATTAGATCTATATCTATCTATCTGTAATCTTT	27000
QY	2701	TGCTGTATTCATTCCTCTTAGAGAGATTTCTATATTAATTCATTCATTAACATTAACGTA	27606
Db	2701	TGCTGTATTCATTCCTCTTAGAGAGATTTCTATATTAATTCATTCATTAACATTAACGTA	27606
QY	2761	TTGCTTAAGTTACCTGCGCAATTCATATAGCTGATGCTATTCGTAATTAAGCTTCATCACTTAA	28202
Db	2761	TTGCTTAAGTTACCTGCGCAATTCATATAGCTGATGCTATTCGTAATTAAGCTTCATCACTTAA	28202
QY	2821	TACATGATATACACAAACCTATTTTTTTTACAATGCTATTCATTAATAATCTACACAAATTC	28808
Db	2821	TACATGATATACACAAACCTATTTTTTTTACAATGCTATTCATTAATAATCTACACAAATTC	28808
QY	2881	CGGTATATCTGTACAAATTCCTACTTAAAGACGGTACTCTAAATGGAAAGACATTTAACT	29404
Db	2881	CGGTATATCTGTACAAATTCCTACTTAAAGACGGTACTCTAAATGGAAAGACATTTAACT	29404
QY	2941	ATAATATTAATCCCTCACAAACCTACACAGCTTTTACTCAGCTTCAATCTTGAAGA	30000
Db	2941	ATAATATTAATCCCTCACAAACCTACACAGCTTTTACTCAGCTTCAATCTTGAAGA	30000
QY	3001	AGACACATTAATTCACATCTATACCTAACAGAAATCTTACTATTTTCCCTATAATTTT	30606
Db	3001	AGACACATTAATTCACATCTATACCTAACAGAAATCTTACTATTTTCCCTATAATTTT	30606
QY	3061	TCCCTCCCTGTATATCTGAGTAATCTTAAATGGGATACGATGCTAAATCTGTAACTTC	31202
Db	3061	TCCCTCCCTGTATATCTGAGTAATCTTAAATGGGATACGATGCTAAATCTGTAACTTC	31202
QY	3121	ATCTATAATAATACGTCACATGATTTGCTGCTCTATAATTCCAATATGAGAGGACTCT	31808
Db	3121	ATCTATAATAATACGTCACATGATTTGCTGCTCTATAATTCCAATATGAGAGGACTCT	31808
QY	3181	ATGAGACAAATTAATTAATACCTTCCCTACACAAATTAATTAACCAAGTAATTAATGAC	32404
Db	3181	ATGAGACAAATTAATTAATACCTTCCCTACACAAATTAATTAACCAAGTAATTAATGAC	32404
QY	3241	TGGTAGCATAGTAGTACATGATATAATAGAGATACATACCCCTTGAATTTTTTATG	33000
Db	3241	TGGTAGCATAGTAGTACATGATATAATAGAGATACATACCCCTTGAATTTTTTATG	33000
QY	3301	TATTAAGCTACCAACTATCTGCTCTTTCCAAACCTGGCAACCAAGTAATGATATACGACT	33606
Db	3301	TATTAAGCTACCAACTATCTGCTCTTTCCAAACCTGGCAACCAAGTAATGATATACGACT	33606
QY	3361	CGATGTAGTACCTGCTTATTAATATCAATCTTCCAAATTCAGTATTAACAGGAGATTAAC	34202

D	3361	CGATGTAGTAGCTGCTTTATTAATCAACTCTCAAAATTAGTAATACAGGGAGTTAC	3420
O	3421	GACTATTTCAATAATCCCTCAAAATGCTGATTTCACTCATCATTTTCTCTACCTAATCTAC	3480
D	3421	GACTATTTCAATAATCCCTCAAAATGCTGATTTCACTCATCATTTTCTCTACCTAATCTAC	3480
O	3481	AGACTATAGCTCTCTTTACAACTAATCTTAATCTTCCCTTGTAAGAAAGGCTTTGTAT	3540
D	3481	AGACTATAGCTCTCTTTACAACTAATCTTAATCTTCCCTTGTAAGAAAGGCTTTGTAT	3540
O	3541	ATAATCAATAGCAACCCATAGCAAGACTTAAACGACGTGGCAATATTAACATGCCACT	3600
D	3541	ATAATCAATAGCAACCCATAGCAAGACTTAAACGACGTGGCAATATTAACATGCCACT	3600
O	3601	AATCATATAAACAAGCAAAATAGGATACCTTTCTTTAAGCTTTTCCAGTACCTTAAATCC	3660
D	3601	AATCATATAAACAAGCAAAATAGGATACCTTTCTTTAAGCTTTTCCAGTACCTTAAATCC	3660
O	3661	ATCAATATCGATCCCTCTTAACCATATATCCAAATAAACAACACAGGCTCTTTTACATA	3720
D	3661	ATCAATATCGATCCCTCTTAACCATATATCCAAATAAACAACACAGGCTCTTTTACATA	3720
O	3721	AGCCATCTTGATCGCGGATAAACCATCAACTGCTAATTTAGTGACATTAATATCATCACT	3780
D	3721	AGCCATCTTGATCGCGGATAAACCATCAACTGCTAATTTAGTGACATTAATATCATCACT	3780
O	3781	TAATATATCTTTTATTTAGATTTCTCGATATCAACTCATCAATCAACATTAATCTTCGA	3840
D	3781	TAATATATCTTTTATTTAGATTTCTCGATATCAACTCATCAACATTAATCTTCGA	3840
O	3841	AATATACAAATCTTTCCTTGACATTTTCAAAATCTGTGCATAATTAACCTACAAATCTAA	3900
D	3841	AATATACAAATCTTTCCTTGACATTTTCAAAATCTGTGCATAATTAACCTACAAATCTAA	3900
O	3901	TGAATATCTGCAACAACAATGATCTCTTTAAATACACCTCAATAACTACGACACACA	3960
D	3901	TGAATATCTGCAACAACAATGATCTCTTTAAATACACCTCAATAACTACGACACACA	3960
O	3961	CTAATGATCAACAACAATATTAATTAATTAATTTTCTTAATTAATTTTAAACATAT	4020
D	3961	CTAATGATCAACAACAATATTAATTAATTAATTTTCTTAATTAATTTTAAACATAT	4020
O	4021	TATATAAATAGTACTATTAATTAACATTAATTAACAGATAGGTTTCTATATATCCAAA	4080
D	4021	TATATAAATAGTACTATTAATTAACATTAATTAACAGATAGGTTTCTATATATCCAAA	4080
O	4081	AAGAAAAATATAAACACTTAATATAAAACCTTATCATCACTTTGTAAACCTATAGTAAT	4140
D	4081	AAGAAAAATATAAACACTTAATATAAAACCTTATCATCACTTTGTAAACCTATAGTAAT	4140
O	4141	AATATAAATATAAATTAATTAATTAATTAATTAACCTCCCTGGAAACAGCATGCTATAGTT	4200
D	4141	AATATAAATATAAATTAATTAATTAATTAATTAACCTCCCTGGAAACAGCATGCTATAGTT	4200
O	4201	TCCATCTAAABACTCATTTGAGCTTTATAGATTTCAATAATTCACGCCAAGCAGAT	4260
D	4201	TCCATCTAAABACTCATTTGAGCTTTATAGATTTCAATAATTCACGCCAAGCAGAT	4260
O	4261	AAGAAATCAATTTGGTATATTTTATATTTTAAACACATAGACAAATATATATCCACCAAT	4320
D	4261	AAGAAATCAATTTGGTATATTTTATATTTTAAACACATAGACAAATATATATCCACCAAT	4320
O	4321	GACTATCAACATATGTATATAGTAGAACCATACCAACAACAAGATATATCATATGACTA	4380
D	4321	GACTATCAACATATGTATATAGTAGAACCATACCAACAACAAGATATATCATATGACTA	4380
O	4381	CTTACCTATATAAATCTAGATAGATTACGACCATCTCATGTATTTGAAGTCAAAATCAT	4440
D	4381	CTTACCTATATAAATCTAGATAGATTACGACCATCTCATGTATTTGAAGTCAAAATCAT	4440
O	4441	CCCGTATATTCATTAATTCACAACAATATATGTCCAAAACATATCAATATATGAATTTTGT	4500

Db	4441	CCGGTATATTCACATTAATCAACAAATAATATGTCCAAAACATCAATGATGAATTTTGT	4560
QY	4501	TATCATATACCAATTTTGGCAAAACAGAAATTTCTTGAACGTTATCTGTAAATTTTG	4560
Db	4501	TATCATATGATACCAATTTGGCAAAAGCAGAAATTTCTTGAACGTTATCTGTAAATTTTG	4560
QY	4561	GTAACAGCCTCTTAACGTAGGCAATTTTGTATGCATGATGATGACTGATTTTGTTC	4620
Db	4561	GTAACAGCCTCTTAACGTAGGCAATTTTGTATGCATGATGATGACTGATTTTGTTC	4620
QY	4621	TCAATACAGCCAAAGGCAAAAAAGATGACCTTGCATCAACAATTACGATTTTCAT	4680
Db	4621	TCAATACAGCCAAAGGCAAAAAAGATGACCTTGCATCAACAATTACGATTTTCAT	4680
QY	4681	CAGAACTACATATTCATTAACCAATCAATCAATATTTTATCAACGTAATGCAATTTGA	4740
Db	4681	CAGAACTACATATTCATTAACCAATCAATCAATATTTTATCAACGTAATGCAATTTGA	4740
QY	4741	CTAATTGATC 4750	
Db	4741	CTAATTGATC 4750	
RESULT 3			
LOCUS	AX685249	1416 bp	DNA linear PAT 29-MAR-2003
DEFINITION	Sequence 56 from Patent WO0206652.		
ACCESSION	AX685249		
VERSION	AX685249.1	GI:29371591	
KEYWORDS			
SOURCE	Erlichia ruminantium (heartwater rickettsia)		
ORGANISM	Erlichia ruminantium		
REFERENCE	1. Barber, A.F., Whitlire, M.M., Kamper, S.M., Simbi, B.H., Ganta, R.R., Moreland, A.L., Wwang, L.D.M., McGuire, T.C., and Mahan, S.M. Erlichia ruminantium polypeptides, antigens, polynucleotides, and methods of use Patent: WO 0206652-A 56 29-AUG-2002; UNIVERSITY OF FLORIDA (US)		
JOURNAL			
FEATURES			
source	1..1416	Location/Qualifiers	
	/organism="Erlichia ruminantium"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:779"		
	1..1416		
misc_feature	/note="Complement to SEQ ID NO:52, nucleotides 2467..3882 Hypothetical nitrogen assimilation regulatory protein Product = 19HMRP3"		
ORIGIN			
Query Match	29.8%; Score 1416; DB 6; Length 1416;		
Best Local Similarity	100.0%; Prid. No. 1.3e-169;		
Matches 1416; Conservative	0; Mismatches 0; Indels 0; Gaps 0;		
QY	2467	TTATTCGAAACATTACACATCCAAAGATTTTAAATTAACGGCTAATGCTGAACGTTC	2526
Db	1	TTATTCGAAACATTACACATCCAAAGATTTTAAATTAACGGCTAATGCTGAACGTTC	60
QY	2527	CATTCCAACAATCAGCAGTTCGTGATACCTTACCTCCAAAAGAGTAACGATTTT	2586
Db	61	CATTCCAACAATCAGCAGTTCGTGATACCTTACCTCCAAAAGAGTAACGATTTT	120
QY	2587	TAAATCTGCTTTCAAATCTTTACGAGCTTTCGTAATGATGACAGAAATTAATCTTAGC	2646
Db	121	TAAATCTGCTTTCAAATCTTTACGAGCTTTCGTAATGATGACAGAAATTAATCTTAGC	180
QY	2647	ACTTAAACATCATTAATAGCGGAGATTAAGATCACTATCTACCTGGTAAATCTTTGCTGT	2706
Db	181	ACTTAAACATCATTAATAGCGGAGATTAAGATCACTATCTACCTGGTAAATCTTTGCTGT	240
QY	2707	AATCATCTCCTAGAGATTTCAATTAATTAATCATCTATTAACATTAGTAATGTCT	2766

D	b	241	AATATATCTCTTAGAGAAATTCATATATTAATAATCCATATTCATATACATTAAGTAATGCTCT	300
Q	y	2767	TAAGTATCCTGGCCATTCAATATGACGTGCATTTGTAATTAAGCTTTCATCACTTAATACATG	2826
D	b	301	TAAGTATCCTGGCCATTCAATATGACGTGCATTTGTAATTAAGCTTTCATCACTTAATACATG	360
Q	y	2827	AGTACCAAAACCAATTTTTTTTACAGATGCAATTCATTAATAATACCTACACAATCCGGTAT	2886
D	b	361	AGTACCAAAACCAATTTTTTTTACAGATGCAATTCATTAATAATACCTACACAATCCGGTAT	420
Q	y	2887	ATCTGTACAAATATTTCTACTAAGAAGCGGTACTCTAATTTGGAAGCACTTTATCTATATA	2946
D	b	421	ATCTGTACAAATATTTCTACTAAGAAGCGGTACTCTAATTTGGAAGCACTTTATCTATATA	480
Q	y	2947	TAAATCTGCAAAACCTACAGCTTTTACTTCACTTCAATATCTTTGGAAAGAC	3006
D	b	481	TAAATCTGCAAAACCTACAGCTTTTACTTCACTTCAATATCTTTGGAAAGAC	540
Q	y	3007	AATATTTCTACATCTTACTATAACAGAAACTTACTATTTTCCCTATATTTTCCCTC	3066
D	b	541	AATATTTCTACATCTTACTATAACAGAAACTTACTATTTTCCCTATATTTTCCCTC	600
Q	y	3067	CTGTATATATCTGAGTATCTTAAATGCGTATGTAATCGTAATCTGTACTTATCTAT	3126
D	b	601	CTGTATATATCTGAGTATCTTAAATGCGTATGTAATCGTAATCTGTACTTATCTAT	660
Q	y	3127	AAATATACGTACCATGATTTTGCTGCTCTATATATCCAAATGAGAGAGTACTATAGA	3186
D	b	661	AAATATACGTACCATGATTTTGCTGCTCTATATATCCAAATGAGAGAGTACTATAGA	720
Q	y	3187	CAATATATTTAATTACTTTCTCACTACCAAAATATTTAACCAAGTAATTAATGCTGTAG	3246
D	b	721	CAATATATTTAATTACTTTCTCACTACCAAAATATTTAACCAAGTAATTAATGCTGTAG	780
Q	y	3247	CATATGATGATACATATGATTTAAATGAGATATACCCCTGGATTTTATATGATATAG	3306
D	b	781	CATATGATGATACATATGATTTAAATGAGATATACCCCTGGATTTTATATGATATAG	840
Q	y	3307	CCTAGCAACTACTTCTTTCCAAACCTGGCGAACCGATATGATATAGACTCGATGT	3366
D	b	841	CCTAGCAACTACTTCTTTCCAAACCTGGCGAACCGATATGATATGATATGACTCGATGT	900
Q	y	3367	AGTATGCTCTTATATATCATCTTCTCAAAATTAACGATTAACAGGGATTAACCGATAT	3426
D	b	901	AGTATGCTCTTATATATCATCTTCTCAAAATTAACGATTAACAGGGATTAACCGATAT	960
Q	y	3427	TTCATATATCTCAAAATGCTATTTCAATCATCATATTTTCTCTAGTATCTACAGATC	3486
D	b	961	TTCATATATCTCAAAATGCTATTTCAATCATCATATTTTCTCTAGTATCTACAGATC	1020
Q	y	3487	TATATGCTCTCTTTACAACTATCTTAACTTCTCTCTGTATAAAGCTTTTCTATATATC	3546
D	b	1021	TATATGCTCTCTTTACAACTATCTTAACTTCTCTCTGTATAAAGCTTTTCTATATATC	1080
Q	y	3547	ATTAGCAACCATATGACAGACTTTACAGAGGCGCAATTAACATGACCCACATATCAT	3606
D	b	1081	ATTAGCAACCATATGACAGACTTTACAGAGGCGCAATTAACATGACCCACATATCAT	1140
Q	y	3607	AATTAACAGCAAAATAAGATATCTTTCTTTTAAGCTTTCCAGTACATTAATCATCAAT	3666
D	b	1141	AATTAACAGCAAAATAAGATATCTTTCTTTTAAGCTTTCCAGTACATTAATCATCAAT	1200
Q	y	3667	ATACAGATCCCTTAATCCATATATCCAAATATACAAATACAGGCTTTTTCATTAAGCAT	3726
D	b	1201	ATACAGATCCCTTAATCCATATATCCAAATATACAAATACAGGCTTTTTCATTAAGCAT	1260
Q	y	3727	CTTGATGCGGATTAACCATCAACTGTCTAATTTAGTACATAATATATCATCACTTAATAT	3786
D	b	1261	CTTGATGCGGATTAACCATCAACTGTCTAATTTAGTACATAATATATCATCACTTAATAT	1320
Q	y	3787	ATCTTTATTAATTTCTGATATCAACTTATCATCATCAACCAATTAATCTTCAGAAATATA	3846

DB 1321 ATCTTTATAGATTTCGATATCACTTCATCAACAACATACTTACGAAATATA 1380

QY 3847 CAATCTTCCTGGACATTTCAAAATCCTGGCAT 3882

DB 1381 CAATCTTCCTGGACATTTCAAAATCCTGGCAT 1416

RESULT 4
AX685247 984 bp DNA linear PAT 29-MAR-2003
LOCUS AX685247
DEFINITION Sequence 54 from Patent WO0206652.
ACCESSION AX685247 GI:29371589
VERSION AX685247.1
KEYWORDS
SOURCE
ORGANISM *Escherichia ruminantium* (heartwater rickettsia)
Escherichia ruminantium
Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
Anaplasmataceae; Ehrlichia.

REFERENCE
AUTHORS Barbet, A.F., Whitmore, W.M., Kamper, S.M., Simbi, B.H., Ganta, R.R.,
Moreland, A.L., Wang, D.M., McGuire, T.C., and Mahan, S.M.
TITLE Ehrlichia ruminantium polypeptides, antigens, polynucleotides, and
methods of use
JOURNAL Patent: WO 0206652-A 54 29-AUG-2002;
UNIVERSITY OF FLORIDA (US)

FEATURES
source location/Qualifiers
1..984
/organism="Ehrlichia ruminantium"
/mol_type="unassigned DNA"
/db_xref="taxon:779"

CDS
1..984
/note="unnamed protein product; Corresponds to SEQ ID
NO:52, nucleotides 632_1615 Hypothetical glycerol
3-phosphate dehydrogenase Product = 19HMOF2"
/codon_start=1
/protein_id="CAD8655.1"
/db_xref="GI:29371590"
/translation="EMTREMBL:CAD8655"
/translation="KKSILGSGSGRTAIALSAGISVNLGRDHNTHTNTYK
NLKPLTHLPNTYATNSIDVLSNNCTILITPTQQLTCTQLOHOMKATP
ILICKSGEITSLEKPESEIAEILQVNTIFLSGSPKAEIHLPGSIVAGNKE
GESLLEISNDVLIKIYQDIIGVIGALNKIATACGIAAGNADNNVAATITG
MWEIKLYIAKHSIDILHTLIGPSCLGLITCTEBSRNAFGLEIGKGRVITLID
HNLKVEGISTVPLISAKUNVELPISIVNLHENISLDAISNLS"

ORIGIN
Query Match 20.7%; Score 984; DB 6; Length 984;
Best Local Similarity 100.0%; Pred. No. 4,46-115;
Matches 984; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 632 ATGAAATCAGATTATTTAGTGCAGATCATTTGGCAGACAGAAATGCAATGCACTGCA 691
DB 1 ATGAAATCAGATTATTTAGTGCAGATCATTTGGCAGACAGAAATGCAATGCACTGCA 60
QY 692 GCAATGCTATTCAGTTACTTATGAGGAGCTGATCTAGAAATATTTCATCAATTAAC 751
DB 61 GCAATGCTATTCAGTTACTTATGAGGAGCTGATCTAGAAATATTTCATCAATTAAC 120
QY 752 ACTTACCGAAAAATTTAAATATTACCCATCATCTTCCAGACAAATATATGCA 811
DB 121 ACTTACCGAAAAATTTAAATATTACCCATCATCTTCCAGACAAATATATGCA 180
QY 812 ACCGCAATATAGCAGAGTATTTATCTGCAACAATATCATGTTATTTCTTAATTCCT 871
DB 181 ACCGCAATATAGCAGAGTATTTATCTGCAACAATATCATGTTATTTCTTAATTCCT 240
QY 872 ACACCAATATGACGACATATGATGACAAATATACAAACAAACAGCATATGTAAT 931
DB 241 ACACCAATATGACGACATATGATGACAAATATACAAACAAACAGCATATGTAAT 300
QY 932 ACTCCAAATTTATTTTGTGTAAGGTATGCAAAATTAATCACTGTAATTTCCAGTGA 991

DB 301 ACTCCAAATTTATTTGTGTAAGGTATGCAAAATTAATCACTGTAATTTCCAGTGA 360

QY 992 ATAGCAGAAATTTTCAATATATCAATTTTATACCTCTGGGCCAGTTTGTCT 1051

DB 361 ATAGCAGAAATTTTCAATATATCAATTTTATACCTCTGGGCCAGTTTGTCT 420

QY 1052 AAAGAAATTCAGAAACATCTCTGTGTATGATGATGCTGTGTATTAAGAACTT 1111

DB 421 AAAGAAATTCAGAAACATCTCTGTGTATGATGATGCTGTGTATTAAGAACTT 480

QY 1112 GGTGAATCTTATATGAAACATTAATGATGATGCTGTGTATTAATTAATCAATCAAT 1171

DB 481 GGTGAATCTTATATGAAACATTAATGATGATGCTGTGTATTAATTAATCAATCAAT 540

QY 1172 ATTATAGGTATCAGATTGGAGCTGCATTAAGAAACATTAATGCAATGATGGAAT 1231

DB 541 ATTATAGGTATCAGATTGGAGCTGCATTAAGAAACATTAATGCAATGATGGAAT 600

QY 1232 ATCGCTGAAAAAATTTAGTATATATGCTGTGTCTACTCTTATTAAGGCAATGAT 1291

DB 601 ATCGCTGAAAAAATTTAGTATATATGCTGTGTCTACTCTTATTAAGGCAATGAT 660

QY 1292 GAAATTAACACATATATATGCAAAAAATCATTCATATGATCTTCATATCAATTAAT 1351

DB 661 GAAATTAACACATATATATGCAAAAAATCATTCATATGATCTTCATATCAATTAAT 720

QY 1352 CCATCATGCTTGGAGATTTATTTATATGATGATGATGATGATGATGATGATGAT 1411

DB 721 CCATCATGCTTGGAGATTTATTTATATGATGATGATGATGATGATGATGATGAT 780

QY 1412 TTGGACTAGAAATGGAAGAAAGTGAATAATATATATATATATATATATATATAT 1471

DB 781 TTGGACTAGAAATGGAAGAAAGTGAATAATATATATATATATATATATATATATAT 840

QY 1472 CTGTGTAAGAACCACTATCTGTAACCACTATATCTTGGCAAAAAATCTTAATGTA 1531

DB 841 CTGTGTAAGAACCACTATCTGTAACCACTATATCTTGGCAAAAAATCTTAATGTA 900

QY 1532 GAATCAATTTGCAATATCTATTTTATCAATTTATGATGATGATGATGATGATGAT 1591

DB 901 GAATCAATTTGCAATATCTATTTTATCAATTTATGATGATGATGATGATGATGAT 960

QY 1592 GCCATTCACCAATATATCTTATG 1615

DB 961 GCCATTCACCAATATATCTTATG 984

RESULT 5
AX685250 597 bp DNA linear PAT 29-MAR-2003
LOCUS AX685250
DEFINITION Sequence 57 from Patent WO0206652.
ACCESSION AX685250
VERSION AX685250.1 GI:29371592
KEYWORDS
SOURCE
ORGANISM *Escherichia ruminantium* (heartwater rickettsia)
Escherichia ruminantium
Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
Anaplasmataceae; Ehrlichia.

REFERENCE
AUTHORS Barbet, A.F., Whitmore, W.M., Kamper, S.M., Simbi, B.H., Ganta, R.R.,
Moreland, A.L., Wang, D.M., McGuire, T.C., and Mahan, S.M.
TITLE Ehrlichia ruminantium polypeptides, antigens, polynucleotides, and
methods of use
JOURNAL Patent: WO 0206652-A 57 29-AUG-2002;
UNIVERSITY OF FLORIDA (US)

FEATURES
source location/Qualifiers
1..597
/organism="Ehrlichia ruminantium"
/mol_type="unassigned DNA"
/db_xref="taxon:779"

misc_feature
1..597
/note="Complement to SEQ ID NO:52, nucleotides 4154_4750
Hypothetical cell division protein ftsO Product =

ORIGIN

19HMOF41"

Query Match 12.6%; Score 597; DB 6; Length 597;
 Best Local Similarity 100.0%; Pred. No. 3.4e-66;
 Matches 597; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4154 TTAATTCATATATATTAACCTTCCGGAACAGCAGCATGTATAGTTTCCATACATAA 4213
 Db 1 TTAATTCATATATATTAACCTTCCGGAACAGCAGCATGTATAGTTTCCATACATAA 60
 QY 4214 CTCACCTGAAGCTTTATAGATATATCAATATATTCAGCCAGCAGCATTAATAT 4273
 Db 61 CTCACCTGAAGCTTTATAGATATATCAATATATTCAGCCAGCAGCATTAATATAT 120
 QY 4274 AGGTATATTTATATTAACCACTAGACAAATAATATCCACCATGACTATCAACTA 4333
 Db 121 AGGTATATTTATATTAACCACTAGACAAATAATATCCACCATGACTATCAACTA 180
 QY 4334 TGTGATAGATGAACCAATACCAACCAAGATATATTCATGACTACTCTCACTATAA 4393
 Db 181 TGTGATAGATGAACCAATACCAACCAAGATATATTCATGACTACTCTCACTATAA 240
 QY 4394 ATCTAATAGATGAAGCAGCAGCATCTCATATGTATGAGTCAATCATCCCATATTA 4453
 Db 241 ATCTAATAGATGAAGCAGCAGCATCTCATATGTATGAGTCAATCATCCCATATTA 300
 QY 4454 ATTATACACATATATATGTCCAAACTATCAATGATAGAAATTTGTATCATGATACA 4513
 Db 301 ATTATACACATATATATGTCCAAACTATCAATGATAGAAATTTGTATCATGATACA 360
 QY 4514 ATTTCGAAAAGCGAATATTTCTGAACTGTATCTGTAAAGATTTGGTAAACGCTCT 4573
 Db 361 ATTTCGAAAAGCGAATATTTCTGAACTGTATCTGTAAAGATTTGGTAAACGCTCT 420
 QY 4574 AACTGAGCATTTTGTATCATGATGACATGACTATTTTGTCTCAATTCAGCCAA 4633
 Db 421 AACTGAGCATTTTGTATCATGATGACATGACTATTTTGTCTCAATTCAGCCAA 480
 QY 4634 AGGCACAAAAGATAGACCTTGATCAACAAGTTAGCTATTTTCATCAGAACTTACA 4693
 Db 481 AGGCACAAAAGATAGACCTTGATCAACAAGTTAGCTATTTTCATCAGAACTTACA 540
 QY 4694 TTGATACCATCATATATTTTATCAACAGAAATTCACAACTGACAAATTTGATC 4750
 Db 541 TTGATACCATCATATATTTTATCAACAGAAATTCACAACTGACAAATTTGATC 597

RESULT 6
 LOCUS AX685246 371 bp DNA linear PAT 29-MAR-2003
 DEFINITION Sequence 53 from Patent WO02066652.
 ACCESSION AX685246
 VERSION AX685246.1 GI:29371588
 KEYWORDS
 SOURCE *Escherichia ruminantium* (heartwater rickettsia)
 ORGANISM *Escherichia ruminantium*
 Bacteria: Proteobacteria; Alphaproteobacteria; Rickettsiales;
 Anaplasmataceae; Ehrlichia.

REFERENCE
 AUTHORS Barbet, A.F., Whitmore, W.M., Kamper, S.W., Simbi, B.H., Ganta, R.R.,
 Moreland, A.L., Mwengi, D.M., McGuire, T.C., and Mahan, S.M.
 TITLE Ehrlichia ruminantium polypeptides, antigens, polynucleotides, and
 methods of use
 JOURNAL Patent: WO 02066652-A 53 29-AUG-2002;
 UNIVERSITY OF FLORIDA (US)

FEATURES
 source
 1..371
 location/Qualifiers

misc_feature
 1..371
 /note="Complement to SEQ ID NO:52, nucleotides <1_371
 /organism="Escherichia ruminantium"
 /mol_type="unassigned DNA"
 /db_xref="taxon:779"

ORIGIN

Hypothetical integrase/recombinase Product = 19HMOF41"

Query Match 7.8%; Score 371; DB 6; Length 371;
 Best Local Similarity 100.0%; Pred. No. 1.3e-37;
 Matches 371; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATCCGCTAATATATTAAGTAAGTTTCTTAAAGATGCGCCAGTAATAAAGACT 60
 Db 1 GATCCGCTAATATATTAAGTAAGTTTCTTAAAGATGCGCCAGTAATAAAGACT 60
 QY 61 GCTTACCTGCTATATATAGTACATACGTATAGATTAACGAAAAAATTTCTCAAGCAG 120
 Db 61 GCTTACCTGCTATATATAGTACATACGTATAGATTAACGAAAAAATTTCTCAAGCAG 120
 QY 121 AGTACCTTGATTTGTATGCTGCTCACTCCACATGTATCTAGAACCAACATTTT 180
 Db 121 AGTACCTTGATTTGTATGCTGCTCACTCCACATGTATCTAGAACCAACATTTT 180
 QY 181 CTCAATCTCCATTTTAACATGACACAAGATTAATTAATGTAACAGAACACCTGTACTC 240
 Db 181 CTCAATCTCCATTTTAACATGACACAAGATTAATTAATGTAACAGAACACCTGTACTC 240
 QY 241 TTATGCAAGAACTTATTAACCTATCAAGGCTCTTACATTAAGAACACTGATTTAA 300
 Db 241 TTATGCAAGAACTTATTAACCTATCAAGGCTCTTACATTAAGAACACTGATTTAA 300
 QY 301 GAATACCTCTCTGCTGTCATCATATTTTCAATATTTTCAATATATTAATATAGAGTCT 360
 Db 301 GAATACCTCTCTGCTGTCATCATATTTTCAATATTTTCAATATATTAATATAGAGTCT 360
 QY 361 TTATTCGACAT 371
 Db 361 TTATTCGACAT 371

RESULT 7
 LOCUS RPKX03/c 279110 bp DNA linear BCT 02-SEP-2002
 DEFINITION Rickettsia prowazekii strain Madrid E, complete genome; segment
 3/4.
 ACCESSION AJ235272 AJ235269
 VERSION AJ235272.1 GI:3861033
 KEYWORDS complete genome.
 SOURCE Rickettsia prowazekii
 ORGANISM Rickettsia prowazekii
 Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
 Rickettsiaceae; Rickettsiae; typhus group.

REFERENCE
 AUTHORS Andersson, S.G., Zomorodipour, A., Andersson, J.O.,
 Sichterich-Ponten, T., Almaraz, U.C., Podowski, R.M., Naslund, A.K.,
 Eriksson, A.S., Winkler, H.H., and Kuland, C.G.
 TITLE The genome sequence of Rickettsia prowazekii and the origin of
 mitochondria
 JOURNAL Nature 396 (6707), 133-140 (1998)
 MEDLINE 98039499
 PUBMED 9823893

REFERENCE
 AUTHORS Andersson, S.G.E.
 TITLE Direct Submission
 JOURNAL Submitted (11-NOV-1998) S.G.E. Andersson,
 Siv. Andersson-Olsson, U. Se, Dept. of Molecular Biology, University
 of Uppsala, Husargatan 3, Uppsala, S-751 24, SWEDEN
 location/Qualifiers

FEATURES
 source
 1..279110
 /organism="Rickettsia prowazekii"
 /mol_type="genomic DNA"
 /strain="Madrid E"

gene
 63..1046
 /db_xref="taxon:782"
 CDS
 63..1046
 /gene="RP479"
 /gene="RP479"

[illegible]

gene
CDS
complement (4969. . 5364)
/gene="RP485"
complement (4969. . 5364)
/gene="RP485"
/codon_start=1
/transl_table=11
/product="NIFU PROTEIN (nifU)"
/protein_id="CA14938.1"
/db_xref="GI:3861038"
/db_xref="SWISS-PROT:Q9ZD61"
/translation="MAISKVLDHEDNPRNATGSLDKKKAVGTGLVGAPACQVMKQ
LEVGDEITDAKPEFGCGSALASSLVETWIKGQVEDAKETKNTLAEILSPV
KLCSLMEIDA KAIADYKQRENKXDS"
complement (5430. . 6662)
/gene="RP486"
complement (5430. . 6662)
/gene="RP486"
/codon_start=1
/transl_table=11
/product="NIFS PROTEIN HOMOLOG (sp11)"
/protein_id="CA14939.1"
/db_xref="GI:3861039"
/db_xref="GOA:Q9ZD60"
/db_xref="SWISS-PROT:Q9ZD60"
/translation="MNQQLKTLTLPIYWDYOSTTPIIDRVMEALPYFTTKFGKPHSR
SHSPREMEKNVENNRSMVAVATGADSEITFGGATSENNLYIKGLAKPFGKKKH
ITLVSEKCVLACRHLHQESIKITLPIKSGNITDLETLMATIDTOLLVVAWANN
ETGVQIPKKEKICRENRVFFSHDIAQGFKLPIVNECNIDLASISGAKTIPKQ
GALVIRPKPRRVATPLINGGQCGMRSGTLPPLIVGLGIASSIAVNEKDKQHVNA
YLFDLFLNYSIKSIEVYINDKQORYGNINLSFAGVEBSIILA-KDLAVSSGAC
TSASLEPPIYVRSIGISEELAHTSIRFGIGRFTTEQEIYAVNLVYCSKIDRLSLSP
MEMQEGVDLKKIRKTAH"
complement (6740. . 7870)
/gene="RP487"
complement (6740. . 7870)
/gene="RP487"
/codon_start=1
/transl_table=11
/product="NIFS PROTEIN HOMOLOG (sp11)"
/protein_id="CA14940.1"
/db_xref="GI:3861040"
/db_xref="GOA:Q9ZD59"
/db_xref="SPTREMBL:Q9ZD59"
/translation="WYLDHNATFTPIDRVKEYIITSLMDKELNPSGATSGRPAAKIT
ETARQKITSLGIMTSREYNTFTSSGTESNMLMKNPYDDITISIEHLSITNNHIN
YAPNKRIRKVRKQGVLDLEHLEDLSQSNASKVLGSVMIANNNGVQIDVAISKIR
KYANFHSIDLVOSEFGKISLNTKEGLDFATISGHKIGGHCQCALINSNFQLPILII
GGGCGSVRSGETENVLALAGGLAEFRRAKQISNRYIKIRLGRLEQKLEYENNAWY
ISNNVRAAGNTLTIVDNTDAOVGLIGEDLNI CVSSGSACSGKISKSHVLTNNQGR
EEEAASSRISLSTHNTSIDKAFIKAFEEIYEKSCNHCI"
complement (7931. . 8125)
/gene="RP488"
complement (7931. . 8125)
/gene="RP488"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="CA14941.1"
/db_xref="GI:3861041"
/db_xref="GOA:Q9ZD58"
/db_xref="SPTREMBL:Q9ZD58"
/translation="WILKIKETIYVAILELATINISNGPIILDKVSIRKINISLNYLQ
IFSRIKSDLVKSIIESKDC"
complement (8277. . 9143)
/gene="RP489"
complement (8277. . 9143)
/gene="RP489"
/codon_start=1
/transl_table=11
/product="unknown"

```

/protein_id="CA14942.1"
/db_xref="GI:3861042"
/db_xref="GOA:Q92D57"
/db_xref="SWISS-PROT:Q92D57"
/translation="MYLSEFLLGLNLTLEQLRLSISIDPYKDYKYNQGYIKYLE
TLEFPIIYICILNITLILKDYNGTSSKVTNTEYIINLPKIKYNSKISVEE
VSEIYLSKNNKIFVIDTGNVSNOKSPVPLENLTKNLIVENTKQHPISIVN
SEI EKNEVILPEIKIIVKYPADNLAPNLFIFGGPVILIFEVTELLERISIVLV
KGLNLTETKTSIOTSRILVMPSSGIPILQPVIIILILFELIILLOLMTFLVFLV
ALLQINKSLSHIT"
complement(9681..10349)
/gene="RP490"
/complement(9681..10349)
/gene="RP490"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="CA14943.1"
/db_xref="GI:3861043"

Query Match      7.5%; Score 357.6; DB 1; Length 279110;
Best Local Similarity 53.1%; Pred. No. 5,9e-37;
Matches 1021; Conservative 0; Mismatches 849; Indels 54; Gaps 10;

QY 2397 TATAATACATCATCTCTTTTAAAGATGACTTCACATTCATACATTAATATTATA 2456
Db 110538 TAAAGAGAAATATAATATGATATGATTTAGCATTTGCTTCTCATATCTTCTT 110479
QY 2457 CATGCATGTTTATTCAGAAACATTCACATCCAGAAATTTTCATTTACGGGTAAAG 2516
Db 110478 CATTTATCTATTTTGCGCGGTATATGAGCTTATATTTTATTTAGTTTCAGTAAAG 110419
QY 2517 CTGAAGCTTCATTCACAAATTCAGAGCTTGATATACATTCCTCCAAAAGAGATA 2576
Db 110418 CCGATCTTCACAAACCAAAATGAGAAAGTTTGAATAATATTATTAACGACTCA 110359
QY 2577 ACTGAGTTTAAATACGTCTTCAAAATCTTCAGAGCTTACGTAATGTCAGAGAA 2636
Db 110358 TTGTGCTGATAGATATGAGCTTCACAAACCTTCTCAGCTTCTTAAGCGGATAGATA 110299
QY 2637 TAACTTAGACATTAACATCATTT-----AATAGCGATATGATATCTACTG 2690
Db 110298 ACATATCAAGCATCTTCAAGTTAGTATGATTAAGACTATTTGCTAATATTTGGAAG 110239
QY 2691 GTAAATCTTTGCTGTAATATCTCTTGA-----GATTCATATTTAAA 2738
Db 110238 GAATCATATAAGTTTATATATTTCAATATTAACCTGTATTAACGATTCATATTAAG 110179
QY 2739 TCCATTCATACATTAAGTATTTGCTTAAGTATCTGCGCATTCATATGACTGATG 2798
Db 110178 TCCATTCACAGGTTACGTAATTTGCTAATATTAACAGCCATTTGATATGAAAGG 110119
QY 2799 CTATTAAGCTTATCTATTAATACATGATACACAAACCTATTTTTCAGAGTCTAT 2858
Db 110118 CTGCAATAGGTTATCGAAAGATGACGTTCTTTAAACCTGAAATTTTGAAGTTGCT 110059
QY 2859 TCAATTAATACCTACACAAATTCGGTATATCTGTACAAATTTCTACAAAGACGTAATC 2918
Db 110058 TACCAATATTTTACATTAAGGTATATCTTCTTCTTCTATTAATGAAGTACTT 109939
QY 2919 TAAATGAGAGCAATTAATCTATTAATATTAATCTCAACAAACCTACAGGCTTTTACTT 2978
Db 109938 TTAAGAAAAAACAATTAAGCGAATATGATCTTCTAGAAATTTCCGTTATTAACCTT 109339
QY 2979 CACTTCATATCTTTGGAAGAGACAAATATCTCACATCTATACATACAGGATCT 3038
Db 109338 CACTTCGATATTTTACGAGTCTGATATATTTTATATCACTTTAGCTTTTTC 109879
QY 3039 TACTATTTCCCTATATATTTTCCCTCTGTATATATCTGAGTAATCTTAATGCGAT 3098
Db 109878 CGAAGAGTTTGA---ATAGTTGATCTTTAAGAAATTTATATATTTTACCTGATAG 109822
QY 3099 CGATGTAATGCTGTAATCTATTAATTAAGCTACATGATTTGCTTCTCTATTA 3158

```

```

Db 109821 GAATAGAAATATATCTGATCTCATCTATATATTAAGTACCTTATTTGCAAACTTA--- 109765
QY 3159 TTCCAAATGAGAGAGTACTCTATAGACAAATATATTATTAATCTTCCCTACATCAATA 3218
Db 109764 -----AGATAGTAGAGAGTTTATATATATCTTCCGCTTTTCCGATTCGCAAT 109715
QY 3219 TATTACCAAGTAATTTATTAAGCTGTAGACATATAGATAGTATCATAGAT ATAAATGAGTA 3277
Db 109714 AATCTTGATTAATTTTTCGTAGATCAGTACAGGTTGAGACTTAAATAAGAAATGATTA 109655
QY 3278 TCATACCCCTTGAGATTTTATATGATATAGCTTACCAACTACTCTTCCCAACCTGGC 3337
Db 109654 TTAAGCTCTTGAATTTGTTTATGAAATTAACCTTGCTGCAAGTTCTTTGCGGCTACGACT 109595
QY 3338 GAACCAATAGATATATGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3397
Db 109594 TTACCATGAATCAATTAATACGATCTGATAGCTGATGATGATGATGATGATGATGATGAT 109535
QY 3398 TTACGATATACAGGAGGATTCAGACTATTTCA---TAATCTCAATGCTGATTTGAC 3454
Db 109534 TTTAAAGTTACGAGACATCTCCATATTAATCAATTTATCTATTAATCTTTGATTTTAA 109475
QY 3455 TCATCATTTTCTGATGATATCTACAGACTCTATAGCTCTTTTACAACTTAAT 3514
Db 109474 TCTATGTTTTCGCTTTTAACTTTGTTACTTCGCAACTCTTTTAAAGTAAATATTAAT 109415
QY 3515 CTGCTCTGTAAGAGGCTTTCTATATATCAATACACCCATACAGAGACTTACA 3574
Db 109414 TTAATATATTAATATGTTTCTCAATATATGATGATGATGATGATGATGATGATGATGAT 109355
QY 3575 GCAGTGCATATTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3634
Db 109354 GCTGTTCTATAGTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 109295
QY 3635 TTAAGCTTTTCCATACATTAATCCATCAATATCAATCTCTTAAACATTTCCAT 3694
Db 109294 TTAATATCTTAAATACCTCAAGCCGATATTTGCTACTTGAACCAATATCAAT 109235
QY 3695 AATCAACATGAGGCTCTTTTCAATAGCATCTTGAATGCGGATTAACATCACTGCT 3754
Db 109234 ATCACTGAGAGACCGGTTTCTCGAAATATTTTAAGAGTTGATGATGATGATGATGATGAT 109175
QY 3755 AATTTAGCATATTAATATATCTCACTTATATATCTTTTATTAATTTCTGATATCACT 3814
Db 109174 ATCTTAGATTTAAACCTTCACTTTTAAATTTGACCAATGATATGCGTATCTTCT 109115
QY 3815 TCATCATCAACATTAATATCTTCAATATATCAATCTTCTGGAATTTCCCAATTC 3874
Db 109114 TCAATGCTATATTAATTAACATCTAATTTGAC---TGACATTTTACCTTAATAATAT 109058
QY 3875 TGTGCAATATTAATCTACATATCTAATGAATATCTGAAACATATGATCTCTTAAAG 3934
Db 109057 AAAAAACCAATGATATGATATGATATGATATGATATGATATGATATGATATGATATGAT 108998
QY 3935 TACACCTATATCTACACACACACATTAATGATCAACCAACATATATTAATTAATCTTA 3994
Db 108997 TTTTGTACAAATTT---TTTGTCAATATATGATGATGATGATGATGATGATGATGATGAT 108940
QY 3995 TTTTACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4054
Db 108939 TATTAGTTGTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 108881
QY 4055 AAGCATAGGTTTCTATATATCCAAAAAGAAATTAATAACCTTAATAAAAACTATC 4114
Db 108880 -----AGCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 108828
QY 4115 ATCACTTTGTAACCTATAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4174
Db 108827 AGTGTGTTTGATCAATAATAATAATAGATATTAATTTAGTGTGTTTAACTGAT 108768
QY 4175 CCC---TGGAACAGGATGCTATAGTTTCCATACAAACCTCACTGAGCTTATATA 4231

```

Db 108767 CACTGATGTAATATAGTATGATATGATACGCAATAAAGCTCTCACTAGTAGTAA 108708
QY 4232 GATATTCATTAATTCACGGCAGCAGCATTAATCATTTATAGTATTTATTTAA 4231
Db 108707 TAAATTTTAAATATGCTTAAACATCTTTTAAATTTATTTGCAATTTATTTGCTTA 108648
QY 4292 ACCA 4295
Db 108647 ACA 108644

RESULT 8
AE008640/c 12477 bp DNA linear BCT 14-SEP-2001
LOCUS Rickettsia conorii Malish 7, section 72 of 114 of the complete genome.
DEFINITION AE008640 AE006914
ACCESSION AE008640.1 GI:15619951
VERSION AE008640.1 GI:15619951
KEYWORDS
SOURCE Rickettsia conorii
ORGANISM Rickettsia conorii
REFERENCE Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales; Rickettsiaceae; Rickettsiae; Rickettsia; spotted fever group.
AUTHORS Ogata,H., Audic,S., Barbe,V., Artiguenave,F., Fournier,P.E., Raoult,D. and Claverie,J.M.
TITLE Selfish DNA in protein-coding genes of Rickettsia
JOURNAL Science 290 (5490), 347-350 (2000)
MEDLINE 20485642
PUBMED 11030655
REFERENCE 2 (sites)
AUTHORS Ogata,H., Audic,S. and Claverie,J.-M.
TITLE Selfish DNA and the origin of genes
JOURNAL Science 291 (5502), 252-253 (2001)
REFERENCE 3 (bases 1 to 12477)
AUTHORS Ogata,H., Audic,S., Renesto-Audiffren,P., Fournier,P.-E., Barbe,V., Samson,D., Roux,V., Cossart,P., Weissenhach,J., Claverie,J.-M. and Raoult,D.
TITLE Mechanisms of evolution in Rickettsia conorii and R. prowazekii
JOURNAL Science 293 (5537), 2093-2098 (2001)
MEDLINE 21442074
PUBMED 11557893
REFERENCE 4 (bases 1 to 12477)
AUTHORS Ogata,H., Audic,S., Renesto-Audiffren,P., Fournier,P.-E., Barbe,V., Samson,D., Roux,V., Cossart,P., Weissenhach,J., Claverie,J.-M. and Raoult,D.
TITLE Direct Submission
JOURNAL Submitted (26-APR-2001) Unite des Rickettsies, Faculte de Medecine, CNRS UMR 6020, Universite de la Mediterranee, 27 boulevard Jean Moulin, Marseille Cedex 05 13385, France
COMMENT A public version of R. conorii genome database is accessible at <http://igs-server.cirs-mrs.fr/>. The database intends to provide updated data. Annotation of the genome is an ongoing task whose goal is to make the genome sequence more useful. Comments to the authors are appreciated.
FEATURES
source location/Qualifiers
1..12477
/organism="Rickettsia conorii"
/mol_type="genomic DNA"
/strain="Malish 7"
/db_xref="taxon:781"
complement(169..267)
/note="REP06, repeated element"
465..636
/note="REP06, A and T rich repeated element"
664..2088
/gene="ntrX"
/note="synonym: RC0849"
664..2088
/gene="ntrX"
/codon_start=1
/transl_table=11
/product="nitrogen assimilation regulatory protein ntrX"

gene
CDS
/protein_id="AA103387.1"
/db_xref="GI:15619952"
/translation="MSQDVLVYDDEBSIRNLIANLKDEGNEPKYANSTOALKILS
KEVSAVAVLDIWLQSGEIDGLIBETIKRYPLEPVIIISGRTLETAVNAIKMAYD
VIEKPPNDKLVILTRACVTRIKRENEDLKYIDKTELVEGCVTLKYMALAKA
AASSCRIMHGKVGSGKELARLHKOSGRVNNPPIIEPTGCTTEKINQIEFGELK
QANKKPTLIEFANGTLYIDEVSNPIPIQVTLKFLKEDOTITKPCGNKIKIDKII
TGTSKTIOEVNNGKLEDELYRINVSILKPSLYEKEDIPLVKPKOISKPSGL
KERSPADETIALQSYEMRGNTROLNVEMLTNPIITGNNELIKPYMISSELIAN
SANITLEDSPDMSMPLEAREBVEROTLSQMSRFNNIKRTSSFMGERSALHRK
LKSLDHPPTNRIINDEYEANA"
2057..3826
/gene="RC0850"
2057..3826
/gene="RC0850"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AA103388.1"
/db_xref="GI:15619953"
/translation="MTKMRKQMLKVIISITTYLLSGSEBSTRDANGHLTDSQSTVI
RNYIISQNKMLKMLKEFGSNLGVKLIGQLINEDLSGIDLSCEITLADPAGSN
LKKAILTNALIOESNPADSVIKIGSHSDPGGSIENHTLQNTNPVGSFDTFENK
TTIINPENSXKSHVMNNNTIDGVNPKALQNNSPKNTITMSIPGCTLECSI
KINFPNNYFESSNLSQTLTIVITKDSNFTOSIREVNFVQSNNSGCFSTAFQDS
TLQNIISLTKCDLQNSTISSVLKHKINNALINNSLNDKNTLSIKNSNPFVIN
KTKGNSITLDNI SYTNPIFSNDPKQFIVINTDLSSEIINNTGQGFNNINFEKSL
INQVNSDVKITLQVNLQVALINSTLTNTAVNSVLSQINNINVOAYSPINNTVS
NNIINSDSSKILPNIVINSGKOLKITHANNLNFQDNLIPRVDSPNSIFK
NMLTNTVTKNSILKEANFSAAILTKTDSNLSILDSIFSKAKIDQAGNNSDLTNAD
PRTALTKSPKAKTSGKYE"
complement(3910..3973)
/note="REP02, RS3-like repeat"
complement(4050..5009)
/gene="RC0851"
complement(4050..5009)
/gene="RC0851"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AA103389.1"
/db_xref="GI:15619954"
/translation="MSKKKQDDQIIKELEERLAKTGTPPSLEEEERPAKIQDPRYI
PKKSGNEVDLILBOLOSLEIKESKSLNDQKQALRDLRLKLEDSPIITGKPID
LSPNDKAITITDIAQADIASVYANRSTEEKKALHLANLEARDIIRIDIPIS
KEYAPKLDIVIASMGKITEGKIKASPAVINIKETIGSSIVNSIKSLPQSSKOOLK
KLVIRYVDLKGIKPKIKERFLKHYERIDQIPRGLIVETAKITQSIIRINGKOKIN
TLKQOQINRMKQSPFALINNSSTPSPFNKAKSSKSNOR"
complement(5047..5136)
/note="REP06, A and T rich repeated element"
5281..5332
/note="REP06, repeated element"
complement(5684..7465)
/gene="pnpA"
/note="synonym: RC0852"
complement(5684..7465)
/gene="pnpA"
/codon_start=1
/transl_table=11
/product="penicillin-binding protein"
/protein_id="AA103390.1"
/db_xref="GI:15619955"
/translation="MANKKILHGEISRPAFIIIGKQGLFSLSLGIRNFYLIISSE
YKTLSDKRNINFLVLPTRGRYLDGDNLTANRKCQYQVLDISINNRRDELEIISN
IINLSEKYNVYIKOKIKSSRTPPIIPDOLMQOVSMIECKHTLASIFIDIGVLR
YFSSVTSHLIGYQINHEKOEHLINISDFNIGSGIKIEYNNKLRGAFGKVE
VNAVQKQVREIAGTPTKSGEDHNLINISLOONIQOYINPKSSAISLVDTAAALVGI
ASTPGFESNHEISLSENYWQSLTSDPYRLINKVIOVNSPPSVKRIITVLAALVGI
NNKTYPCDGSALGTNSRCNHSQGHGLDMGSLKSCNLYWEILARIVGPDKILE
VAREFQGSKTGIDLAPSSSGVPSKEMKKKUKLPSIGDSFNALIGCGFAGVTM
LARFTALASNGKLTPTRLKNDPFPVNNIKIPNTIIQSLVNTVAVAGTAYNR
ILGKROLAGKGTQVOGKLNADKDLRSKSIAMRRHNAFLGAPYHDPBSVTVF
IDHGGGGSVAAAPVARKIMSVDLDKYL"

gene complement(7455..7976)
CDS /gene="RC0853"
complement(7455..7976)
/gene="RC0853"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAL03391.1"
/db_xref="GI:15619956"
/translation="MKIKQEIYKNSLOIILKLPALICPTIIFPVNSYKINQSR
PMAEILFIYFMSLQNIISIFLGLFDIGISGPIGMSLVFLSAVITYKLSKY
FAKNYLINFIIICVYCLFIIINFKYLIITIKLEADGYLIIFQELTIIFSYNIIRLV
LDSPMYFKKYK"

repeat_region 8041..8068
/note="REP03, repeated element"
complement(8226..8420)
/gene="RC0854"
complement(8226..8420)
/gene="RC0854"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAL03392.1"
/db_xref="GI:15619957"
/translation="MSKNKNTDKPSIFIPWIKSFDLTTTNNITSKQLFISTSEAO
EELISDNDKRNPSCELL"
complement(8595..10280)
/gene="pdp2"
/note="synonym: RC0855"
complement(8595..10280)
/gene="pdp2"

CDS /codon_start=1
/transl_table=11
/product="penicillin-binding protein"
/protein_id="AAL03393.1"
/db_xref="GI:15619958"
/translation="MKOIMKWKPKIKSLITIMVCSKTKRILMTVCPSFLFECTIS
VRLIVATSVYKINNSFKKQKQPKKXIVDRGNLILANESASLIPANPOIVDPETS
VAKAEILPDIKAKLAELKSNKSFIVKRDVLPSPQEKITISGLGFEFEBOCKI
YTESLILKLVGVYGRDVGSLGELAYDKYLTNSDEVLPKPKKEPLQSIDIRQA
SLSEIDIKTLQKFAIGAVGIADPNNGELIALVKNEDPDPYPSLAKPEELNINS
LGIVMGSVFKALIMAVGFDIGVIMNDAVIDISVAKVGFGKDYTPROGHSVPEF
LYSNIGSOIKLEIGKSNPKYKLGGLDLOLBEPREGTPLPSEKRNNEELSVT
MSYGISISPLHFYAMLPAVNGGTLVDLLIKRQEKVIGTVESENSTOKKIF
BAVVEGNGKPAKGVKYLIGKTKTAECISGAGGKKYLNKNSAASFLENLPSNQ
YIIFRDPBPKTKESFPRATNSWTAFTAGVFERMISLYLEPIESES"

repeat_region complement(10325..10355)
/note="REP02, RS3-1-like repeat"
10418..10443
/note="REP02, RS3-1-like repeat"
complement(10447..10472)
/note="REP02, RS3-1-like repeat"
10511..10634
/note="REP04, repeated element"
complement(10624..11022)
/gene="RC0856"
complement(10624..11022)
/gene="RC0856"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAL03394.1"
/db_xref="GI:15619959"
/translation="MSIRKLHYTLITLIIICSLFSIKERVSTLDYQLSVVKQINS
ENNINIHIAEKAVYLIPARLEKLAALYKLETVKSYQMIDPLGNIDONIKENHI
SISKSKRYKTIINCKICITVSSKVDKTR"
complement(11029..11952)

gene
Query Match 7.1%; Score 338.6; DB 1; Length 12477;
Best Local Similarity 55.7%; Pred. No. 4.5e-34;
Matches 793; Conservative 0; Mismatches 589; Indels 41; Gaps 6;
QY 2437 TTCATACATAATAATATATACATGATGTATTATCAGAAACATTACCAATCAAGAT 2496

Db 2079 TGCCTTCATATATCTTCGTCATTTATCTATTGTAGAGGATATGTAGCTTATATA 2020
QY 2497 TTTCATTTACGGTGTATGCTGAAGCTTCATTCACAACAAATTCAGAGTTGTGTATAC 2556
Db 2019 TTTTAATTTAGATGCAAGCCGATCTTTCATACCGACAAATGAGAGATGTTTGGAAAT 1960
QY 2557 ATTTACCTCCAAACGAGATACGAGTTTAAATACGTCTTCAATTTCTTACAGAGC 2616
Db 1959 ATTTAATTTAAACGAGCTCTTGTGCGCATATATAGACGCTCAAAACCTTCTAGC 1900
QY 2617 TTTACGATAGTACAGAAATACCTTAGACCTAAACATCATTT-----AATAGGCA 2670
Db 1899 TTCTCAGGCGCATGATATACATCAACGATCTTCAAGTTTGTATAGATTACACCT 1840
QY 2671 ATTTGATATCTATATCTACTGTGTAATCTTTGCTGTATCATCTCTTAGA----- 2722
Db 1839 ATTTGCTAATATTTTCAGAAAGTATCATATAGGTTTAAATTTTATTTACTCTAGT 1780
QY 2723 ----GATTTCATATTAATAATCCATTCATTAACATTAAGTAATGCTTAAAGTACCTG 2778
Db 1779 TAAGGATTCATATTAATTAAGTCCATTCACACAGCTTACGTAATGTTCTAATATGCCC 1720
QY 2779 CCAATCATATGACTGCAATGCTATTAAGCTTCATCATTAATACATGAGTACACAAC 2838
Db 1719 CCAATCATATGAGTTGAAGACGAGCTAGTTTCATCGCAAAAGCTACGTTCTTTAAAC 1660
QY 2839 TATTTTTCACAGATGCTATTCATAAATACCTACACAAATTCGATATCTGTACATA 2898
Db 1659 TGAATATTTGAAGTGTGCTTAACAAATATCTTAAGAGTGGGATATCTTCTTTCT 1600
QY 2899 TTCTACTAAAGCGTACTCTTAATTTGAAGACATTTAATCTAATATTAATCTCTACA 2958
Db 1599 TTCTATTAATGAAGAGTACTTTTGAAGAGTACATTAAGCGCATATATAGATCTCTAG 1540
QY 2959 AAACCCACAGCTTTACTTCTTCACTTCATTAATCTTGAAGAAGACACAAATATCTCAC 3018
Db 1539 GAATTTGCCATTTATACCTTCACTTGATTTTGAAGATTCGGGTATATTTTAT 1480
QY 3019 ATCTATACATACAGGAATCTTACTATTTTCCCTATATATTTTCCCTCGTAATATCT 3078
Db 1479 ATCAATTTAATATTTTTCGCCGAAGTTTGA---ATAGTTGATCTTAAAGAAATTT 1423
QY 3079 GATTAATCTTAATTTGGTATCGATCTGTAATCTGTAATCTTCACTATTAATTAAGTAC 3138
Db 1422 TAATTAATTTTACTTGATAGGAATAGGAATATTAAGACTTCACTATATTAATTAAGTAC 1363
QY 3139 ATGATTTGCTTGTCTATATATTCATATAGAGAGTACTATAGACAAATATATAT 3198
Db 1362 GTATTTGGCAATTTCA-----AGATATAGAGGGGTATATATTTGCGCT 1319
QY 3199 ACTTTTCTACATCAAAATATATTAACAGTAATTTAGCTGAGATAGATAGTA 3258
Db 1318 GCTTTTCAATTCGCCGAATATTTCTTGATGATTTTTCGAGTACATACAGTATGGGC 1259
QY 3259 CATAGAT-ATTAATGAGATATACACCTTGATTTTATGTAATTAAGCTACAGACTA 3317
Db 1258 TGAATATATGAGAGATATTAACCTTTTATGTTGTTATGATTAATATCTTGGCGCA 1199
QY 3318 CTTCCTTTCAACACTGGCGAAGCATATAGATATAGATCTCGATGATGATGCTT 3377
Db 1198 GTTCCTTTACCGTACCGACTTTACCGTGAATCATTAATAGACAGCTAGAGGATGCTCT 1139
QY 3378 TATTAATCATCTTCAATTAAGTATACGATTAAGAGGAGTATCCGATATTTCA---TAAT 3434
Db 1138 TTGCTATTCGATTTATATTTTAAGTATACGACATTCGCTATATTCAGTTTAT 1079
QY 3435 CCGTAATGCTGATTTCAATCATCATCTTCTTCAAGTATCTACAGACTATATGCTC 3494
Db 1078 CTATAACCTTTGATTTAATATATATTTTCCGTTTAACTTTGTACTTACAAAGCTC 1019
QY 3495 TCTTTCACTACTTATCTCTCTGTAAGAGCTTCTATATCATATAGACAC 3554

Db 1018 TTGTAAGTAATACTAATTAATCATATTAACGGTTTCTATATATCGTAAGCAC 959

QY 3555 CCATATGACAGACTTATAGACAGCGCAATTAACATGCCCATATATCAATTAACAG 3614

Db 958 CCAATTTTATGATTAATCTGCTGTTCTATAGTACCGGACCGCTAATTAATATATACG 899

QY 3615 GCATTAAGAGAACCTTTCTTAAGCTTTTCCAGTACACTTAATCATCAATATGACATC 3674

Db 898 GCATTAAGAGAACCTTTTATATATCTTAATCAATCCAGTCGCTCAATTTGCTTC 839

QY 3675 CTTTAAACCATATATGCAATTAATACAACTCAGGCTTTTTCATAGCAATCTTGATCG 3734

Db 838 CTTGAAGCCAAATATATAGTACAACTGACAGACCGGTTTTCGAAAGTATTTAAAGG 779

QY 3735 CGCATTAACCATCACTGCTAATTAATTAAGTACATTAATTAATCACTTAATATCTTTA 3794

Db 778 CTTGAGTACTATAGGCGACCTTAGGCTTAACCTCATCTTCAATTTGACAGCA 719

QY 3795 TTGATTTCTGATTAATCACTTCACTGACACAACTAATCTTC 3837

Db 718 TGAATTTCTGATTAATCACTTCACTGACGCTCACTAATTAACATC 676

RESULT 9
AF176314 23906 bp DNA linear BCT 07-SEP-1999
LOCUS Zymomonas mobilis fosmid clone 42b11, complete sequence.
DEFINITION AF176314
ACCESSION AF176314 GI:5834286
VERSION
KEYWORDS
SOURCE
ORGANISM
Zymomonas mobilis
Zymomonas mobilis
Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
Sphingomonadaceae; Zymomonas.
REFERENCE
AUTHORS Lee, H.J. and Kang, H.S.
TITLE Sequence analysis of 42b11 fosmid clone of Zymomonas mobilis ZM4
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 23906)
Lee, H.J. and Kang, H.S.
DIRECT SUBMISSION
SUBMITTED (07-AUG-1999) Microbiology, Seoul National University,
JOURNAL Sanse-1 Shillimdong Kwanakgu, Seoul 151-742, Korea
LOCATION/Qualifiers
1. 23906
/organism="Zymomonas mobilis"
/mol_type="genomic DNA"
/strain="ZM4"
/db_xref="taxon:542"
/clone="fosmid 42b11"
104. 4699
/gene="gltB"
104. 4699
/gene="gltB"
/note="glutamate synthase large subunit"
/codon_start=1
/transl_table=11
/product="GltB"
/protein_id="AAD53891.1"
/db_xref="GI:5834287"
/translation="MSIKSSADSPFVFAERELAKGMYRPFEPENDACGVGLVAATDQ
KSRVVTALIDALKAVMHGRVADAGDGGAGHNDLPASFPDDALAAHGAFLPN
BLAVGVEPRPRDLSAQETCRITVSEIIDRGRIYGRQYPIVSIYGEKATRE
ISQITAGHPEDMSANFENKYLIRRIEIKVIAQTNKISARCLVRSIVKGL
IAEALPTCYQDERFVSRAVFFQRSTVFPKWMIAQTNKISARCLVRSIVKGL
NMKSHKELKMAISLAFENSEDKPLIPAGASDTALDAVEFTICRAGHRAPAALMI
PEAMQHPMPFAHMDYSYLAIVMEPMGDPAALITDGSNAVAGVDNRNALPRTVR
TDNLIIVSESGGVVVENNYIAKQGLPGOMAVNIDEGRFEDRLKLIASEBP
YGMITKDFKSLDLKSYDPIPYDRDELIKROVASGLTLEMDIVISPAEFGKA
MSGMDPTPLVLTETKPRITSGFQNFQNSQVTPNIDSLRETQVMALTRRGNTL
DSENRKQILITIDPYLKTSENQGLKDFGEQAVIICTFSPDGAALRLASNKT
QENAVRSGHSEFLTDEHISDRVAVMILATAVHTHLIRKELRSASIVNVAISC
IDPHVAVALIGVATTVAAYLAESAIAVRHSGGLFGTWLTLEQCLDNHHALICGLIKI

MAKGIJAVSVSYGAVNPEAVGLBALVNDLFPQMTAKISGEGYASLORSALLRHONA
PANKKALDVCGEYRORASGDSHAAYTAELMRLOKXVSQSDVTTYFERKEHOLDPV
YLDLILGFMSKOPPIEBEVEPTTEIRRRPFIPTGMLGLSPAHETTLAMNRITAK
SVSGEGEAPERHPYENGDNANSPIKOYASGRFQVTAAYLQACEIEIKVQKARFG
EGGLSPGFVFTDMLAKLRSTPGVTLISPPPHDLYSIEDLQIYDLKQINPNARVG
VKLSAGSIGITIAAGVAKAHADSIMISGPNVGTGASPLTSIXYAGTPMKGLSEANOV
LTLNGLRHVKTARDGLKTRGDIVAJIIGAEFEGIGTSLIAMGLVROCHSNKC
PVIGCTODRAMOKFIPTEPERVINYMTFAEYREILAKGVRSIKEIIGOTDLISQV
NCGADLDDLDLNPULAKVDANTEARVPOIDEPFNVPSILDAOMLQDAAYFTREK
TOLTYSRVNDRALGTRFSALTRNGQMSLADGVHVPTRGSAQSLGAFSVKGLTL
EVRGDANDYVKGSLGATITVPTVSSPIDGHTNTITGTYVIGATAGCLPAAGRGE
RFARVNSGAEVTEECGANGCYMTGLVALIGSVGPNFAGMTGMAVLDEDSFE
RHANPNIIIMORLASHMEVTLRSMIADAVTDSFWARVILDEWDLNFKWQICPK
EMGLRLEVPDLDKNTEDQKEKAKGPELKEQEGAF"
4779. 5450
/note="hypothetical protein; ORF1"
/codon_start=1
/evidence="not experimental"
/transl_table=11
/product="unknown"
/protein_id="AAD53892.1"
/db_xref="GI:5834288"
/translation="MMQLYQPLCPFSKAVRFLGKGIYELIRESPLKREDFVDM
NPAGCTPVWVDSAPVILVDSOVCIFEEETDRMPLISGTAVRAEIRLVSLEPNK
LYPEYTAIPMNERMLKRLVHRQAPDAIIRGMKXANSLSHLDYIMLIDRLRAGVTL
SLADFAPAAQIVSDYLGVDWRGHNREMYAAMKSRSPFPIIAHREVIILPPDHY
DQDPF"
complement(6178. 6519)
/gene="hima"
complement(6178. 6519)
/gene="hima"
/codon_start=1
/transl_table=11
/product="Hima"
/protein_id="AAD53893.1"
/db_xref="GI:5834289"
/translation="NKKRYKDNRTNGSQMOADIPTLITRADITMLYHEVGLSPADSAXM
IEOMLGHITDALKKGENVKISGGSFILDKNERVORBNKGTIEVPIABRVLTFRAS
OLMORRIKGA"
complement(7332. 8759)
/gene="ntrX"
complement(7332. 8759)
/gene="ntrX"
/note="nitrogen assimilation regulatory protein"
/codon_start=1
/transl_table=11
/product="NtrX"
/protein_id="AAD53894.1"
/db_xref="GI:5834290"
/translation="NALREILVVDENDIRELVAGVLDDEGVATRTAASDATALQALAE
RPSPLVLDVWLKSGLDLGLLELEBIKREDPPLPVIVISGHNIDINVAIAIQGAVD
IEKPFASERLLILVERATETERLEKRNALKEITELTGNSSAINVAATLAKVA
GTGSRVLTITGPAVGKAVARLLHMSPPAAAPFIYSAAMPEPVEGEELGVESSQ
GVYRPLLEKAGGTLFLDEIDAMPPTQAKTILVTDOSFMRVGEERVAVKIDRVYS
ATKARDAREIAGRRREDIYXRLNVVPHLPALABRERIPITLYVAVHYATPERSVA
TFQVSEALALQADIDPGRVQKRVNERTIILAPGRLSLIDVMLFSEVTLGQNS
LTADGQSTRSIIITPLREARETFERYLRVQIRRSQNSIRASFTGERSALHRLKX
SIGLAEPNANATSNNDNNENEG"
complement(8759. 11065)
/gene="ntrY"
complement(8759. 11065)
/gene="ntrY"
/note="nitrogen regulation protein"
/codon_start=1
/transl_table=11
/product="NtrY"
/protein_id="AAD53895.1"
/db_xref="GI:5834291"
/translation="MMEDSALISTEKTEVDRKKIRLMPLSKRLYYFLRRPPANET
LTFVAVLVAIFSYRLTYHGTPAPRPSLYVAILLYNTVYAMMLMTLARIARLR
AAKSPIDGGGRGHRVRLVAFSTVAAPFILLVTFASLIFQYVSWFSPSARITLANA
DRVAOSYVENYEDRIIGDINAMPNDFRALSGAPLEDPMWSFTIQOLINRLNEVAL
LQMPEDITCAPIMLAGLITDRPLENRLPAKMIALLRKQATSSISSGRLIEAAILLD

ASRLTYIYRQDLPVLLQASFAKTALEDYRNLVESSVLIQIRFNGALLVLSLLIVA
IAIYAFVADVVRPISHNVEAVRIGSDLGARTHLKHNRDEIGTLGFANRANK
LEEQTLMLVNDLSDRAFAVRISAGVLSVNSQHTIQLVLSQKLLKXDEK
ILIGELDLNAPLSDILAQAGDSIVHILVNDPEPTFAVRVADNNVLFEDDITO
QLADORAAMADVARIIAHEIKNPLTPIQLAARLORRYGROITSDEFTSPRTSTIV
ROYGDLRMYDESSFAKMPKPYRKEALISDIGHRLPLHEAHAPQIGVFPAQDLP
LWCDOROLGRALTNVKNAVEAIEENDRGSGHWMTIKREDSDHILVADTGILP
KESRLTEPTWTRVGTGLATVKRIEELHAGTLTFSDRGGGTIVALSFDLVALA
PVLISVAEGKIGNOEPALPGLTRNSE"
/complement (11516..12550)
/gene="nifH3"
/codon_start=1
/transl_table=11
/product="NifH3-like protein"
/protein_id="AAD53896.1"
/db_xref="GI:5834292"
/translation="MNSPEHPTNHLISPIQSVRIIDNPVILAMTGVTLDPFRIV
RBYGSLANTEMIASQAMIRETROSLOKSTWDPSEPTSVQVLCCTPDPMESALNQ
DRGAALIDIMGCPYKRYVSGDAGSLKMDLKLASIIETATKASLPTTLKRGMD
LNSLNAPETARIKODIGVQWTVHGKTRQLIKGHADVAFVQVDAVSIPVAGDI
NSLDAKTALQOQSGADGIVNGRSGYKPMILQMQGLDGMKAPSIDBYRITTE
HYDMLSYGTLITGVNARKHIGWYTKGLPASAEFRANKINREDEDDRVKRNAGRILPT
LAQPNRLTD"
12702..13865
/note="hypothetical protein; ORF2"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="unknown"
/protein_id="AAD53897.1"
/db_xref="GI:5834293"
/translation="MNRKQTNALIVAAQGVVQKALOKOYRKTAKRANLAHARDN
LAPRIDIYOVYVADGHQTLYOBAYGDRPLPPVYIGVTRRVSIVNGLKAAHDKRYR
VLIHDAARPLKTVLIDRLDLAKSKKALPYLPVADTLVNDVEAVVRNMFHRQYR
QAFDELTVIAHQAMTGSDEPTDQAVVAFKELALVAGDLLEKLTYPDEFVAEA
OMTEKMSVCGGDFVHCPEAGDHMLGKIKIPHDHGLKADADVALHALTDALGA
IADGIDGTFPSPDPMKGNANSTOFLFVAVALAKKAGAIIDHADVTVCAPKVPYR
PAWRKILPREFMGKNGSASKORQLKSLVLAQAKGSLPKRLPVFACRIFCAN"
13873..14412
/note="hypothetical protein; ORF3"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="unknown"
/protein_id="AAD53898.1"
/db_xref="GI:5834294"
/translation="MNPVIRHSIALVACVGTGMAAATAEAGSOSQCVTHAEQALM
ETRLPDLQAKQCSVLDSDSYLKKQATLSKFRQSESTDMKAKAEPRLGSA
VADMGAGRAALITWVSDIVSQKMPQSCETVNGFLPVAEOPANKIGDVSVFAL
LKKIGVGRALPFDICSD"
14327..15103
/note="hypothetical protein; ORF4"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="unknown"
/protein_id="AAD53899.1"
/db_xref="GI:5834295"
/translation="MFLPSPCLRLKRLADVLCILLFVREIKPWTMDNNDGDSQ
PADIKEALQAEQODISADSEKLOKTIQNMENQOENVNHLPELVTASLVI
DLNRAAGRIIAVESCTGGLVMAALTEVPGASDVDAQFVYANAKIDILNIGDVI
EERGSIAVAVMAAAMNAVEKSDADIAVLTIGAPTGDERKPGVTVPARARAD
PHEVVAEOKSPFDLGRSGIRLQALCALSLMPDSISQ"
/complement (15290..15733)
/note="hypothetical protein; ORF5"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="unknown"
/protein_id="AAD53900.1"
/db_xref="GI:5834296"

Query Match 4.9%; Score 231; DB 1; Length 23906;
Best Local Similarity 49.9%; Pred. No. 1.3e-20;
Matches 693; Conservative 0; Mismatches 660; Indels 36; Gaps 3;
CDS
gene
/translation="MPEYTESKILPYTPQQLPDLVADISHYPERLPVIAVRIRREE
NMRATDIFVGRKARESTSKVLDSHSVSEVITDPLSHLNMHFTFEEBGRTRL
DFVDFRSPRIEPLALGQFDDAVQKCTQAFERRANNIKGLSDA"
/complement (15733..16704)
/gene="lipA"
/complement (15733..16704)
Qy 2489 CCAGAAATTTTCATTTACGGTGTAAATGCTGAACGTTCCATTCACAAATTCAGCGTT 2548
Db 7402 CCAAGGCTTTTAACTTACATGTAGAGCCGAGCGCTCCATACCAATTAAGAGCTGTT 7461
2549 CGTATACATTTACCTCCAAACGAGATACGATGTTTAAATCTGCTTTCAATTCCT 2608
7462 CGAAGATATTTGCTGAGAAACGCGCAATTTGCACCCGAGATATTCCTGTAAGGTT 7521
Qy 2609 TCACGAGCTTTACGTAATGCTACAGAAATTAATTAGCACTTAAACATCTAATATGCG 2668
Db 7522 TCGGTCCTTCTTAAATGGGTGCGCATATACAGCGGCTGATTCGCGTCGCGTC 7581
2669 GAATTAGTACTATATCTACGTAATCTTTGCTGTAAATCATCTCCTTAGA----- 2722
Db 7582 AAGCTGTTTATGATGCCAGCAGCTCGAAGGAGCATATCGACATATTCGGCTTAA 7641
2723 -----GATTCATATTAATTAATTCATTTATTAACATTAAGTAAATGCTTAAGTTA 2773
Db 7642 CGATGCCCGGTGCAAAATTAATGTCGTTCAACCATTAAGAGCTGTCGACCTTA 2701
Qy 2774 CCGGACCATTCATATGACGATGCTATTAAGGCTTCATCACTTAATACATAGTACAC 2833
Db 7702 CCGGACCATTCATATGACGATGCTATTAAGGCTTCATCACTTAATACATAGTACAC 2761
2834 AAACCTATTTTTCATAGATGCTATTAATTAATTAATTAATTAATTAATTAATTA 2893
Db 7762 ACAGAGGCTTCTGTGATATGAGCCAGCAAAATAGCGATGAGCATGGAATATCTTCG 7821
2894 CAATATTCATCTAAGAGCGGATCTAATTAAGGAGCATTTATCATTAATTAATTC 2953
Qy 7822 CGGCGATCGCCAAAGCCGTAATGACCGGAAACGACATTTAGCGATTAATTAATCT 7881
Db 2954 TCACAAACCTTACAGCTTTTACCTCACTTCAATATCTTTGAGAGAGACACATTAAT 3013
7882 TCACGAAACCGGCTTACGAGATTCGCGTAAGTCCGCGCGGAGAGACACAA 7941
Qy 3014 CTCACATCTATACGAAAGAAATCTTAATTTTCCCTATATATTTTCCCTCCTGAAT 3073
Db 7942 CGGCGATCGACCTTACCATCTTTCGCGACCGACCGCATTAAGCTCTGATCAGTCAGA 8001
3074 AATTCGATATCTTAATTCGATCGATCGATATCGTAATCTGTTATCTTCAATTAATTAAC 3133
Db 8002 ACTGTAATAATCTTGCGCTGTGCGAATGGGATATTCGCGATTCATCAAGAAACAGT 8061
Qy 3134 GTACACGATTTGCTGCTTATTAATTCGAAATGAGAGGTACTCTATGAGCAATATA 3193
Db 8062 GTTCTTCATGAGCTTTTCAATTAAGCCGGGCGA-----ACCACA 8103
3194 TTATTTACTTTCCTACATCAACCAATATTAATTAACCAATTAATTAATTAATTAATTA 3253
Db 8104 CTTGATATCTTTCATCAACCAATTAATTTCTTCGCGCCGCTGCTGATTCATCAATGCGC 8163
Qy 3254 GAGTACATTAATTAATGAGATCAATACCTTGATTTTATATGATTAATTAATTAATTA 3313
Db 8164 GCCGAAACGACATTAAGGCGGAGCGGACGAGCATCCACATTAAGAGCAATTCGGGCG 8223
3314 ACTACTTCCTTTCACACCTGCGAACCAGTATGATGATAGATAGATAGATAGTATGATGAT 3373
Db 8224 GCGACCTCTTTCCGAGCCGAGAGGCGCTGTTATTAAGAGCGGCGTCCGCTGCGA 8283
Qy 3374 GCTTATTAATCACTTTCGAAATTAAGTAAACAGGAGGATTAACGACTATTT---TCA 3430

Db	8284	ACCGTTTCAACGAGCCGACGATGATGCACTGGAAATGCCGTTAATCTTCA	8343	/strain="56601"
Qy	3431	TAACTCCAAATGCTGATTTCAACTATCTATTTCTCTACGTAATACCAAGCTCTATA	3490	/serovar="lai"
Db	8344	GTAAGCCGATTTTGGCCTTTAAAGAGCGTTTCTGCCGCAATCTTTCGTTGGGT	8403	/db_xref="taxon:189518"
Qy	3491	GCTCTCTTACAACTAATTAACTTCTCTCTGTAAGAGCGTTTCTATATATCATPA	3550	/chromosome="1"
Db	8404	GCTCTTTCGACAAACATATGTAACCGTTCGCTTCAAGAGGTTTCTCGATGAATCAAG	8463	/note="Strain 56601 is maintained by the Institute of Epidemiology and Microbiology, Chinese Academy of Preventive Medicine"
Qy	3551	GCACCCAAATGACAGACTTATACAGAGTGGCAATATTACATGCCCATATATCATATA	3610	/locus_tag="LA2397"
Db	8464	GCACTTGTGCGAATAGCCGTTACGCCGTTCAATTTGCAATGCGGAAATACCAAGC	8523	/codon_start=1
Qy	3611	ACAGCAAAATGAGATACCTTTCTTAAAGCTTTTCAGTACATTAAATCATCATATCA	3670	/transl_table=11
Db	8524	ACCGCAAAAGTGGATACCGCCGTTGATTTCTTCCAAATATTCAGACCGCTCCAGTTTC	8583	/product="hypothetical protein"
Qy	3671	GATCCTCTTAACCATATATCATATATACATACATGCGTCTTTTCATTAAGCATCTTG	3730	/protein_id="AA049596.1"
Db	8584	GAACTTTTACCAAGATCATCAATTAACCAAGAGCGCTTTCTCTTAAAGCTGT	8643	/db_xref="GI:24196157"
Qy	3731	ATCGCGAATTAACCATTAATCTGTAATTATGACATTAATATCATCACTTAATATCT	3790	/translation="MLGSKASDFHTALPKYILSSQNKLAGPPIEL"
Db	8644	AAATCTCCATCATCTGTCGAGCGCTTCTTGTGCAATACCTTCATCATCAAGAACCCCC	8703	/gene="pria"
Qy	3791	TTTATGATTTGATATCATATCATATCAACAATTAATCTTCAAAATATATCAT	3850	/locus_tag="LA2398"
Db	8704	GCTACCAATTCAGAAATGTCATTTGCTATCATCAACCAAGAAATTTTACGCCCATATT	8763	/complement(328..2271)
Qy	3851	CTTTCCTTG 3859		/gene="pria"
Db	8764	CGCTCCTTG 8772		/locus_tag="LA2398"
RESULT 10				
AE011408				
LOCUS	AE011408	14211 bp	DNA	linear
DEFINITION	Leptospira interrogans serovar lai str. 56601 chromosome I, section 217 of 397 of the complete sequence.			
ACCESSION	AE011408	AE010300		
VERSION	AE011408.1	GI:24196156		
KEYWORDS				
SOURCE	Leptospira interrogans serovar lai str. 56601			
ORGANISM	Bacteria; Spirochaetales; Spirochaetaceae; Leptospiraceae; Leptospira.			
REFERENCE	1 (bases 1 to 14211)			
AUTHORS	Ren,S.X., Fu,G., Jiang,X.G., Zeng,R., Miao,Y.G., Xu,H., Zhang,Y.X., Xiong,H., Lu,L.F., Jiang,H.Q., Jia,J., Tu,Y.F., Jiang,J.X., Gu,W.Y., Zhang,Y.Q., Cai,Z., Sheng,H.H., Yin,H.F., Zhang,Y., Zhu,G.F., Wan,M., Huang,H.L., Qian,Z., Wang,S.Y., Ma,W., Yao,Z.J., Shen,Y., Qiang,B.Q., Xia,Q.C., Guo,X.K., Danchin,A., Saito, Hiroshi,I., Somerville,R.L., Wen,Y.M., Shi,M.H., Chen,Z., Xu,J.G. and Zhao,G.P.			
TITLE	Unique physiological and pathogenic features of Leptospira interrogans revealed by whole-genome sequencing			
JOURNAL	Nature 422 (6934), 888-893 (2003)			
PUBMED	12712204			
REFERENCE	2 (bases 1 to 14211)			
AUTHORS	Ren,S., Fu,G., Jiang,X., Zeng,R., Xiong,H., Lu,L., Lu,G., Jiang,H., Ding,Y., Jia,J., Tu,Y., Gu,W., Cai,Z., Sheng,H., Yin,H., Zhang,Y., Zhu,G., Wang,S., Shen,Y., Qiang,B., Chen,Z., Wen,Y., Xu,J. and Zhao,G.			
TITLE	Direct Submission			
JOURNAL	Submitted (12-MAR-2002) Chinese National Human Genome Center at Shanghai, 250 Bi Bo Road, Shanghai 201203, China			
COMMENT	Updated information will be available at our World Wide Web site (http://www.chng.sh.cn/lep/). Comments to the authors are appreciated.			
FEATURES	Location/Qualifiers			
source	1..14211			
	/organism="Leptospira interrogans serovar lai str. 56601"			
	/mol_type="genomic DNA"			

gene
/strain="Brutus"
/specific_host="adzuki bean beetle, Callosobruchus chinensis strain jc"
/db_xref="taxon:188349"
complement(1.. .2857)
/gene="pol"
/pseudo
complement(1.. .2857)
/gene="pol"
/pseudo

CDS
/codon_start=1
/transl_table=1
complement(3049.. .4212)
/gene="ORF1"
complement(3049.. .4212)
/gene="ORF1"
/codon_start=1
/transl_table=1
/protein_id="BAC22720.1"
/db_xref="GI:24475508"
/translation="MDCPPRGAIPSTTGSAGVALONTPTASGSEKAEIRPFLIV
KKKSDFKVNPPLISRTMYGLIDGVYKIKSTNDGILVERSSKQARQLNOIQFPAPF
EVEVAHPTLTKISKGVITCSDLNCNIEICNELSSIGVLEVRIRKSRQMLVDITAN
HILFNKPTLKEIKVANMTLKRPYIISPPLRCNQCKFGHTTRCSFOKICVGGKQPF
HEGTCDSPPICTPCQGNHPAOSKQCIIKYEEPAIQOLKVKEISTYEAKRNVAVOYPP
TPNVYSAAASTLTLDQILLALKTIIIOETINREVOOLRPVPPOFTTAORTSLSLP
PFRPRDTKCSSEKRKRSSTLPGPSEEAASSDSSATKTKKGWPKXPRK"
complement(4899.. .6097)
/gene="ftsZ"

gene
/gene="ftsZ"
/pseudo
complement(4899.. .6097)
/gene="ftsZ"
/pseudo

CDS
/codon_start=1
/transl_table=1
/product="cell division protein ftsZ"
6097.. .6495
/gene="BMEI0172"
/pseudo
6097.. .6495
/gene="BMEI0172"

CDS
/codon_start=1
/transl_table=1
/product="cell division protein ftsZ"
6097.. .6495
/gene="BMEI0172"
/pseudo
6097.. .6495
/gene="BMEI0172"

```
gene      /transl_table=1  
          complement(6489. .6955)  
          /gene="RP741"  
          /pseudo  
CDS      complement(6489. .6955)  
          /gene="RP741"  
          /pseudo  
          /codon_start=1  
          /transl_table=1  
gene      7043. .7835  
          /gene="sdhB"  
          /pseudo  
          7043. .7835  
          /gene="sdhA"  
          /pseudo  
CDS      /gene="sdhA"  
          /pseudo  
          /codon_start=1  
          /transl_table=1  
          /product="succinate dehydrogenase iron-sulfur subunit"  
          complement(8318. .8731)  
          /pseudo  
          /codon_start=1  
          /transl_table=1  
          /product="putative acetyltransferase"  
          complement(9071. .9469)  
          /gene="pgpa"  
          /pseudo  
gene      /pseudo  
          complement(9071. .9469)  
          /gene="pgpa"  
          /pseudo  
          /codon_start=1
```


Query Match	4.4%;	Score 207.8;	DB 6;	Length 349980;
Best Local Similarity	45.4%;	Pred. No. 4.3e-18;		
Matches 123;	Conservative	0;	Mismatches 147;	Indels 37; Gaps 12

[illegible]

QY 2106 AATTACCAAAAAAGTTAAATATACATCTCTCCCAAAATATGAGAGATACATACAC 2165
 Db 213626 AAAAT 213567
 QY 2166 GTCAAAAT 2225
 Db 213566 ATAAAT 213507
 QY 2226 TAT 2285
 Db 213506 TAT 213447
 QY 2286 ATGAT 2345
 Db 213446 AT 213387
 QY 2346 TAACAAT 2405
 Db 213386 TAT 213327
 QY 2406 TCATATCTTTTAAAGATGACTTCACATTCATATATATATATATATATATATATAT 2464
 Db 213326 ATAAAT 213267
 QY 2465 GTTATATCAAAACATTCACATTCACATTCATATATATATATATATATATATATAT 2524
 Db 213266 ATAAAT 213207
 QY 2525 TCCAT-TCCACAAATTCACAGCTTCGTATATATATATATATATATATATATATAT 2583
 Db 213206 CAT 213147
 QY 2584 TTTTAAATATCTGCTTCAAAATCTTCACAGCTTCGTATATATATATATATATATAT 2643
 Db 213146 TAT 213087
 QY 2644 AGCACTTAAACATTCAT 2703
 Db 213086 AAT 213031
 QY 2704 TGTAT 2763
 Db 213030 ATAAACAT 212971
 QY 2764 TCTTAAATATCTGCTTCATATATATATATATATATATATATATATATATATATAT 2823
 Db 212970 ACTTAT 212911
 QY 2824 ATGAT 2840
 Db 212910 AAAACAAAAAAACCAA 212894

RESULT 13
 AB063522/c 349174 bp DNA linear ECT 22-NOV-2002
 LOCUS AB063522
 DEFINITION *Wigglesworthia brevipalpis* DNA, complete genome, section 2/2.
 ACCESSION AB063522 BA000021
 VERSION AB063522.2 GI:25166251
 KEYWORDS
 SOURCE *Wigglesworthia glosiniidia* endosymbiont of *Glossina brevipalpis*
 ORGANISM *Wigglesworthia glosiniidia* endosymbiont of *Glossina brevipalpis*
 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 Enterobacteriaceae; *Wigglesworthia*.
 REFERENCE
 AUTHORS Akman, L., Yamashita, A., Watanabe, H., Oshima, K., Shiba, T.,
 Hattori, M. and Akao, S.
 TITLE Genome sequence of the endocellular obligate symbiont of tsetse
 flies, *Wigglesworthia glosiniidia*
 JOURNAL Nature Genetics 3 September 2002; doi:10.1038/ng986;
 REMARK <http://www.nature.com/cgi-taf/DynaPage.taf?file=/ng/Journal/vaop/ncurrent/full/ng986.html>

REFERENCE 2 (bases 1 to 349174)
 AUTHORS Yamashita, A., Oshima, K., Furuya, K., Yoshino, C., Shiba, T. and
 Hattori, M.
 TITLE Direct Submission
 JOURNAL Submitted (21-JUN-2001) Masahira Hattori, Kitasato University
 School of Science, Kitasato 1-15-1, Sagamihara, Kanagawa 228-8555,
 Japan (E-mail: hattori@genome.sci.kitasato-u.ac.jp),
 Tel:81-42-778-9926, Fax:81-42-778-9926
 COMMENT On Nov 21, 2002 this sequence version replaced gi:19071886.
 FEATURES
 source
 location/Qualifiers
 1..349174
 /organism="Wigglesworthia glosiniidia endosymbiont of
 Glossina brevipalpis"
 /mol_type="genomic DNA"
 /specific_host="Glossina brevipalpis"
 /db_xref="taxon:36870"
 138..404
 /gene="rpsT"
 138..404
 /gene="rpsT"
 /note="rpsT"
 /note="Protein synthesis"
 /codon_start=1
 /transl_table=1
 /protein_id="BAC24442.1"
 /db_xref="GI:25166252"
 /translation="MATIKSKKRSIOSEKRRKNSKSVIRSFIKVNSILKKEK
 DTINSEIEMQSLIDRYAKNLHKNSASRLSKSLKFKIR"
 complement (424..1614)
 /gene="nhaA"
 complement (424..1614)
 /gene="nhaA"
 /note="Transport and binding proteins"
 /codon_start=1
 /transl_table=1
 /protein_id="BAC24443.1"
 /db_xref="GI:25166253"
 /translation="MKKFKCKSKFPKFPFSGLLILFCFLAIFISNTNFKTYNI
 INYPLIKINHDKFSLTNIIVNDIMTFPELEIGIEIKHEMLVGLSKNSRALPGIA
 AIGMIFPALIVNTRKEDSSIGMAITATDIAFVAVGLKILGHSIPSLILFELS
 LAIFDDGALLIATFYSNHIIDQYMLISLVILITLSINYLAIVTCIYIYIGILIM
 ESTPLSGHISIGVLDIGLIMPSSTYSSKSKSMARLKSPLKRLITLTFAPF
 NSGINFSENFENLSSILPFGIFGLVAGKIGVLSYSVKRSLKLPGLISFKEIA
 GISFLCIGIFENLISFLSNLAFQNIENKIIYIAKFSILISISVSVIGFLVLYLKYKI
 KLNK"
 complement (1675..2799)
 /gene="dnaJ"
 complement (1675..2799)
 /gene="dnaJ"
 /note="Protein fate"
 /codon_start=1
 /transl_table=1
 /protein_id="BAC24444.1"
 /db_xref="GI:25166254"
 /translation="MAKSDYVDLIGISKNSAPREIKTAYKRLAVKPEHPNPNLEAE
 SKREIKAEYEVLLDPKRAATNQGHTAFDQSSNTERFSNADPSDIDEDVGDIFG
 NNRKRSNRSGSDIQYININLSLEAVAGISKEITIPLERNTISGKMPGTSIAQNS
 TCGOGQIQMRQGFSEVQOCPSGRCKIRFPCKNSGOGVRESKINSIRIPAGV
 DNDKDIRLSEGBAGRLGSGSDLVYKISVKEHPFKREBNHLYCEVPINFAMAALGG
 DIWVPLADGVKRLIPAEOTGRLFPMRGVYSISGHCGLLCRVVETPVRLSEK
 OKSLKEIGISFGRGTOMNSPSSKSPFGVKKFFPGLN"
 complement (2874..4811)
 /gene="dnaK"
 complement (2874..4811)
 /gene="dnaK"
 /note="Protein fate"
 /codon_start=1
 /transl_table=1
 /protein_id="BAC24445.1"
 /db_xref="GI:25166255"
 /translation="WEILKNGITIGDLTNSGVALIENGKXVYINSBGRTPSII
 AYENELIIVGQPAKQSVNPNTEFALRLIGRFTHEVORDVINIPYKIVSEN
 GDVWLWVKQKVAEPQISAEILKMKKTEADYIGKSLTBAVITVPAVFNDTQCAKDK
 AGRIAGLDVRLIINEPTAALAYGLDKTKGNRIIYAVDGGGTFDISIIEIDVDGEK

gene

CDS

TFEVLSTNGDTHLGEDEPDRLLINLYVNEFKKEQIDLRNDPLAMORLSESEKAKIE
LSVYQTVNLPYITADSSGPRNNIKYTRAKLSLVLELYKTLEPVKSLDQAKK
LIDKIVLGGQTMPLVOKKSDPEPKERKQVNPBEAVAIGAAGVGVLAQKAK
VLLDVTPLSGIETMGVMTLLSKNTTTPKSCIPSTAEQNSAVTHLYOGEEK
RSIDPKSIGQRNLDIAPAMGMPQIETTFIDADGIHRSKAKRSGREKTTAKS
SGLSNEIDKMLKESANAABDIKFEELVKQNPADHLHSTRKQILEANPLERKT
EIKCINLELSIKGEDKDIKIQSLQISGLVPSKNKENKEDIKIKNKEN
NKNTDVPVDAEFEBIKDKKN"
gene . 6391
/gene="fabA"
5858 . 6391
/gene="fabA"
/note="Fatty acid and phospholipid metabolism"
/codon_start=1
/transl_table=1
/protein_id="BAC24446.1"
/db_xref="GI:25166256"
/translation="MVKKKFYEKEDLLCSRGELPGKYGQPLPAPYMLINDLRVKT
ENGWYNGEFLKALDIPNMPFSCHPIDPVWPGCLGDAPWQVFLHMGAG
KGRALGRVKEVFGSGLPTSKIVYTHFRITNKLFGWADGEVFCDEKIIYTAAD
LKVGLFDITSFKDFK"
gene . 7077
/gene="yhgI"
complement (7084 . 7707)
/gene="yhgI"
/note="Unknown function"
/codon_start=1
/transl_table=1
/protein_id="BAC24447.1"
/db_xref="GI:25166257"
/translation="MFKYIINFYEYKLLPYKMMIISKAKHYLSKLSKKK
GNTKLSIENKDKNTLCQIKYFPDKKEENYEKFNLPDEYVYKNIPIYIKKID
IESKIDKEIILIKI EKRENLKRNISNITKYNKSLPYGGAPEYKIDENNFVILK
FEGGNGCSMAKVLKESIEKEIKKNPNKNGVIDITDIHSESSF"
gene . 7803
/gene="pro-TGG"
complement (7730 . 7803)
/gene="pro-TGG"
complement (7876 . 7950)
/gene="his-GTG"
complement (7876 . 7950)
/gene="his-GTG"
complement (7960 . 8034)
/gene="arg-CCG"
complement (7960 . 8034)
/gene="arg-CCG"
complement (8250 . 9398)
/gene="weeC"
complement (8250 . 9398)
/gene="weeC"
/note="Cell envelope"
/codon_start=1
/transl_table=1
/protein_id="BAC24448.1"
/db_xref="GI:25166258"
/translation="MKKKNKIIPNLPYSGKEIKYIDPAIRNKTSGSGYPSNLCTL
WIKNLNTKVFLLPSCTSSLEMAALLKINKDEIIMPSTFVSTNPNVLGAKIV
QVDIRSDNLINENLIESAITKTKAIVVHVAGISCEMDFIKNISRYKLWIEDSA
QSMASKYKLYLGLIGHGCFSPHETKRYSGSGGAILLINDELLKRAIIEKGN
RSLFKIKTKYVWADIGSSYTLIDIOAYVMSQKISEKIOKREMLNVPISLEN
IKLPIIPFNKCKNANATFYLLPNDKTCLPIEYMKNKITLLRHYVVLHSPAGK
KRGKRGNDITKTSNNLVRLPLFPLNKNKNDQXIIISNIYPIDKXS"
gene . 10677
/gene="rho"
complement (9418 . 10677)
/gene="rho"
/note="Transcription"
/codon_start=1
/transl_table=1
/protein_id="BAC24449.1"
/db_xref="GI:25166259"
/translation="MNTRELKNIKLSIDILITLGENVGLNLANRKKODITSTIKQAK
IGEDI FGDGVLEILQDEGFLRSSDSSSLAAGPDDIYVSPQIRRFNLRITDITSGIKR

PEKEGRYPALLKVDENVYDKPSSRNKILFENLTPLHANSRLKMERGSTEDLTTR
VLDIAPIGORGGLIIVAPKAGKMLLONIAQSI SNYHPOCVLLVLLIDEPBEVTE
MORIVKGIATSTPESSRHVQVSEMT EKKRLVEHKQVITLLDSTTRAAVNT
VYPAGKVLITGVDAANALHRPREFGAARNEEGSLITIALLLDTSKNDVYEE
FPGTGMELHLSRKIAEKRVFPADYNVSGTRKEBLTSSBELQMMILIRIHPMNE
IDGMEFLWKLSEVPKINIKESIGYKLVKSIPSPFLPFKEKIDKXIGIVSQKLE
KISEKYVNLSEVPKINIKESIGYKLVKSIPSPFLPFKEKIDKXIGIVSQKLE
IRIDENLRL"
gene . 12788
/gene="pgk"
11604 . 12788
/gene="pgk"
/note="Energy metabolism"
/codon_start=1
/transl_table=1
/protein_id="BAC24450.1"
/db_xref="GI:25166260"
/translation="MNNKEIISVNDDEFEINVLKQKRPVLYDFMAQMPCKSMIPIL
KISEKYVNLSEVPKINIKESIGYKLVKSIPSPFLPFKEKIDKXIGIVSQKLE
IRIDENLRL"
gene . 12788
/gene="pgk"
11604 . 12788
/gene="pgk"
/note="Energy metabolism"
/codon_start=1
/transl_table=1
/protein_id="BAC24451.1"
/db_xref="GI:25166261"
/translation="MKIKVTDNLKDKKVLIRSDLVPIKNGIKMCMRIKSLIETI
NVIKNEKACVIVASHLGNPIEEKDYNPSLHIVNKKIKYERLIONLNEFA
KEKEIIVENYFNGSRKNDKLSKRYANCDIIVMSFGSSHSESTGIIKFPAP
ISEIGLFLKEIKYKALPNSKRIYVIFGSKISTGLVKEKSNSENVLVGGAI
ANTLIPSGHNGKSLHDKNNISYKIKELVKNIIIPDPIYTDQINSGYSKESIK
EIKNDPIYDIKNSCECFIKIKAKKATPNNKPGLEIIOFNGTEINSVHSS
AKSITGGETVLAHMFNFIIDRFSTISGGAFSLERKELPVLSLKKRKYTNFK
QK"
gene . 13881
12811 . 13881
Query Match 4.3%; Score 204; DB 1; Length 349174;
Best Local Similarity 47.8%; Pred. No. 1.3e-17;
Matches 828; Conservative 0; Mismatches 880; Indels 26; Gaps 7;
QY 400 ATAGTACTATCTAATTAATAAATCTCAATGCTTTTGCAAGAACTACAAATATA 459
276130 ATTGTTATATATTTTAAAAAATAATATATATATATATATATATATATATAT 276071
DB 460 ATGAGCACTTAATTAATATATATATATATATATATATATATATATATATAT 519
QY 276070 AATTAAT 276011
DB 520 TAATATACATCTGACATGATTAACAAATGTAACACATTTCTTAACCTAGATAATAC 579
276010 TATTTAATTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 275951
DB 580 AATTAATTAATTAATTAATTTTACGCGTTCTTGATATTAATTAAGTAAAAATGAAT 639
275950 AATTAATTAATTAATTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 275895
QY 640 CAGTATTTAGGTGACAGATCAATTTGGACACAGCAATAGCAATGCTGACGACATGG 699
275894 TACTGTAATTTGAGCGTGGAGCTTATGAGCAACGCTGCGTGTTCCTTTTCCAAAAAAA 275835
DB 700 TATATACGTAATCTATAGGCGACGATCATATGAATTTTACATATPAAACACTTACCG 759
275834 TAGAAAAGTGTTTTATAGGGGATGAAATATAAACAATAGCATCTCTTAAAAAAGATCG 275775
DB 760 AAAAAATTTAATATTTTACCACATATCATCTACAGACACATATATGCAACGACAA 819
275774 TTGCAATATAAATTTTATCCAAAATTAATTTCTTAATGATTTAAAAATGAAATTTTC 275715
QY 820 TATGACGAGATATATCTGACACATATCATGTTATTTTAACTATCTATCTACACACAA 879
275714 TTTAAAAAAGCTATTAATAATATATGCAATATC---TATGTAATAGTGTTCCTAGATATAG 275658

QY 880 ATTACGACCATATGTACACAAATACACAAAGCATATGTGTAATACTCCAT 939
 Db 275657 TTTTAAAAACATTTTAAATTC-----AACATCTTAAATTAAGATGTTTTT 275604
 QY 940 ATTAATTTGTAGTAAAGGTATGTGAATATGATCATCTCAATTTCCAGTAAATGCGA 999
 Db 275603 AATATGTGAACCTAAAGGGTTGAACCAAGAACAGAACATATTAACAAGAGTGTGTA 275544
 QY 1000 AGAATTTTACATATATATCAATTTT---TATCTCTGCTGCTCAAGTTTGTCAAGA 1056
 Db 275543 CGATATCTTAGGAAGGAAGAACGTTTGCAATTAATACAGCACTCTTTGCAAGTA 275484
 QY 1057 AATGGAACACATCTTCTTGATATATGATATGATCTGCTGCTGTA TAAAGAACTTGTA 1116
 Db 275483 TCTAGTATGTGATTAACCACTCAATGCTACTACTATATATTTGTAATAAATCTTGA 275424
 QY 1117 ATCATGTATAGAAACATATATGATATGATCTTCTAAATAATATATACATCAAGATATAT 1176
 Db 275423 ATTTTATAGCAATTAATCAATTAATTAATTTTAAATATATACAGATTTCTGATTTAGT 275364
 QY 1177 AGGTGACATGATGAGCTGCACTTAAAGAACATTAATGCAATGCAATGGAATATGCG 1236
 Db 275363 TGGAGTGCAAAATAGGAGAGATATAAATATGTCGATAGCTTCTGATATGCTGA 275304
 QY 1237 TGGAAAAATTTAGGTAATATGCTGTTGCTACTGTTATTAATAAGCAATGAATGAAT 1236
 Db 275303 TGGAAATGCAATGGTCTTCAAGCTTAAACAGCTTTATTAATTTGCGATTAAGAAAT 275244
 QY 1297 TAAACACATATATATAGCAAAATATCATTAATGATCTTCAATTAATTTGCTCATC 1356
 Db 275243 GATATA-----AGTGGGTAAAGTATATGAGGAGCAATGATCACTTTATGAGGAATGTC 275190
 QY 1357 ATGCTTGGAGATCTAATATTAATTAACATGTAACAACAACATTCACCAATATGCTTTGG 1416
 Db 275189 TGGAGTAGAGACCTTGTTCTTCAACATGATGATATGATATGAAGAACAGAAATTTGG 275130
 QY 1417 ACTAGAAATAGAAAAAGTGAATATATATATTAATAGATCAACAACCTAAAGCTTGT 1476
 Db 275129 GATATATAGCTCAAGATATAGCATAGAAATCTAAATCAAAAGTGGATGATATAT 275070
 QY 1477 TGAAGAACCACTATCTGTAAACCACTGATATCATTTAGCAAAAAAATTAATGTAAGT 1536
 Db 275069 AAGAGGATACATATATATTAAGAGATTTTATTTTATCATGTAACAACAATTAACAT 275010
 QY 1537 ACCAATTTGCTATATCAATTTATTAATATGAGAAATATATCATAGATTAAGCAT 1596
 Db 275009 GCTTAAATCAACAAGTTTAAATTTTAAATATGATATGATCAACAGTAAATAATCAT 274950
 QY 1597 ATCAACATATATCTTATCTATCATCTTGTGATTAATTTCAACAATGATTAATTAAC 1656
 Db 274949 TTTAATATTTTAAACAAATAATTAATTAATTAATTAATTTTAAATTTTAAATTA 274890
 QY 1657 GGTCAAAAAT--ATGAATACAGCTGACCAATATTAATAAAGCATCAATACCAT 1724
 Db 274889 AATTAATATGAAATATTTAATTAATTAATTTAATTTGAGTTTACTTTTAAAT 274830
 QY 1715 CGTTAATCACTAATGATGATATATTAATTAATTAATTAATTAATTAATTAATTAAT 1774
 Db 274829 AATATAGTTTACCTTAAATAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 274770
 QY 1775 TTCCTAGCTAATTAATAAAGAAATTCATCTCTATCTCAATTAACCAATGTTATCA 1834
 Db 274769 AATTAATTAATTAATTAATTAATTAATTAATTAATTTTAAATTAATTTTAAATTA 274710
 QY 1835 GATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1894
 Db 274709 AAGCAAAATTTTAAATCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 274650
 QY 1895 TTTCACATCATCATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1952
 Db 274649 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 274590
 QY 1953 ACCAATACATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2012

Db 274589 TTAATAATTAATTAATTTTAAATTTTAAATTAATTAATTTTAAATAATTTT 274530
 QY 2013 CTGACATTAATTAATTTGATATATCTTGCATATTTTCAACATTAATTAATTAATTT 2072
 Db 274529 TAACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 274470
 QY 2073 TTTTATGCCATTAATTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2126
 Db 274469 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 274416

RESULT 14
 PVGNABC/c 4362 bp DNA linear BCT 17-FEB-1993
 LOCUS P.vulgaris genes glnA, ntrB and ntrC.
 DEFINITION X68129
 ACCESSION X68129.1 GI:45899
 VERSION X68129.1 GI:45899
 KEYWORDS glnA gene; glutamate--ammonia ligase; nitrogen regulator; ntrB gene; ntrC gene; two-component regulatory system.
 SOURCE Proteus vulgaris
 ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Proteus.
 REFERENCE 1 (bases 1 to 4362) Seglitz-Morsdorf, U., Morsdorf, G. and Kaltwasser, H. Cloning, heterologous expression, and sequencing of the Proteus vulgaris glnA ntrC operon and implications of nitrogen control on heterologous urease expression
 JOURNAL FEMS Microbiol. Lett. 106 (2), 157-164 (1993)
 MEDLINE 93202435
 PUBMED 8095910
 REFERENCE 2 (bases 1 to 4362) Morsdorf, G. Submitted (17-AUG-1992) G. Morsdorf, Universitaet Bayreuth, Lehrstuhl f. Mikrobiologie, Universitaets str.30, 8580 Bayreuth, FRG

FEATURES
 source location/Qualifiers
 1..4362
 /organism="Proteus vulgaris"
 /mol_type="genomic DNA"
 /strain="DSM 30118"
 /db_xref="taxon:585"
 37..51
 /bound_molecy="NR I"
 71..85
 /bound_molecy="NR I"
 160..175
 /note="-24/-12 signal"
 259..1668
 /gene="glnA"
 259..1668
 /gene="glnA"
 /BC_number="6.3.1.2"
 /codon_start=1
 /evidence=experimental
 /transl_table=11
 /product="glutamate--ammonia ligase"
 /protein_id="CAA48234.1"
 /db_xref="GI:45900"
 /db_xref="GOA:P28786"
 /db_xref="SWISS-PROT:P28786"
 /translation="MSLEHVLSTIEBKRYVLDLPTDTRGKQHLTPAHQVNDPE
 ESKKFDGSSIGWKGINESDVLMPDSTALDFPQDPLIISCDVLEPQMGYD
 RDPKEDSKRAEDYLKSSGIADTVLPSPSEDFLPDIRKNDISGASVAINDIEAAM
 TPTKEDKNGKRPWKGGYPLPVDSSODIRSTNCNIMBMGAVVEAHHEVATAG
 ONEVATRPVMTTKKADETQIYKVYQNVAVHVGKATFMPKPLVGVNGSGMCHMSLS
 KNGTULPADDKYGK.SEMALVYIGGIKGAALNPTNPTNSYKRIYGVGEPAPVMA
 YKARNSASIRIPVYASMKYARIEVRPDLANPYLAPACIMAGIDIGITNTHGDA
 MDKLTLDPEEKKEIPTYAGSLBEALNLDADREFLTHGVFTNDSIDAYLALRTD
 VQRVRAHPHELEMYSA"
 1675..1707

-35_signal

```

terminator      1735..1752
protein_bind    1824..1838
                /bound_molecule="NR 1"
gene            1854..2900
                /gene="ntcB"
                /length="2900"
                /map="ntcB"
                /codon_start=1
                /transl_table=11
                /product="NR 11"
                /protein_id="CAA48235.1"
                /db_xref="GI:45901"
                /db_xref="GOA:P28788"
                /db_xref="SWISS-PROT:P28788"
                /translation="METGNLPDNTLLIDSLHSVLVINEPFITCYANHAALOYLQAGSR
                RKLEPTPTDLSYHSFPAELMOETLARGQSFIDNEVILVTHNQSHMSLSAQISEQ
                HILIELAPMDSORLSQEOIQAOQIARELVGLAHEIKNPGLGAGLAQLKSLP
                DPALTEYTOVILBOADRRLTLVDRLLGPQPKTKTGS IHVVERVAQLISLECPENV
                TLKDYDPSLSELSHPDOI EOVILNTIRNALOAVERTGTTILIRTRTAQITLHGER
                HRVARIIDVIDTSGSIPPHLODILFYPMWGREDDNGSLISARINVDHAGKLEPETS
                MPNTFESITPIK"
                2911..4332
                /gene="ntcC"
                2911..4332
                /gene="ntcC"
                /codon_start=1
                /transl_table=11
                /evidence=experimental
                /product="NR 1"
                /protein_id="CAA48236.1"
                /db_xref="GI:45902"
                /db_xref="GOA:P28787"
                /db_xref="SWISS-PROT:P28787"
                /translation="MOKGNVWVDDSSIRFWLERAITTBGMLCRAFEHANVDLKALN
                SEQDVLISDIRMPDMGLSLIKTIKQVETPLVITMTASDIDAAVAVOCAGAPKYL
                PKPPIDIBETLALIRATTHREKQKPNANENLOSVDITGAPMAQEVETRIIGLSR
                SSISVLINGESGTEKELVAFALHRSFRALPFTALMNAIIPDLLESELPHEKFAF
                TGAQVQGRERQANGSLFLDELIDMPDLIDITRLRLVLEGGQFVGVGAPVKAQVLR
                IIAATHODLERKRVESGFRDLVHRLNVIRIQLPPLADRTEDIPLARVFLQTALEK
                GVETKSLHESGSLKTMMEVYVSGNVROLCRMILTWTWTAQOIMPODLPEILSLAEK
                AKINIRLTSQMSQHSLSLMADEALGEEKENIINDALPQFERTLLSLAVTQHKMDA
                ARLLGMRNTITRLKELGIEDY"
ORIGIN
Query Match      4.0%; Score 188.6; DB 1; Length 4362;
Best Local Similarity 49.6%; Pred. No. 5.4e-15;
Matches 575; Conservative 0; Mismatches 569; Indels 15; Gaps 3;

```

```

QY 2699 TTGGCTGAATCCTCTAGAGATTTCATTAATCCATTCATTAACATTCGT 2758
DB 4052 TCTTGTGGCATGATTTCTGGCTTGACGTCAATACCGTCAGCCACACACATTTTCG 3993
QY 2759 AATTGCTTAACTACCTGACATTCATGATGATGATGATTAAGCTTCACTT 2818
DB 3992 AGTGTCTTACGTTCCAGACCAACGATTCATGATGATGATTAAGCTTGTATGC 3933
QY 2819 AATACATGATGACCAAACTATTTTTCACAGATGCTATTCATTAATACCTACACAA 2878
DB 3932 AGAGATTTGGTTTCAACACCCAGCTCTTTGGCCGTTTTCAGAAAAGTAGCCGCTAAA 3873
QY 2879 TCCGATATATCTGATACATATTTCTACTAAGACGCTACTAATTTGAGACATTTAAT 2938
DB 3872 CTAGAAATATCTTCTGTTCTATCCGGAAGGTGTGATTAAGATGACAAATCACTTTAG 3813
QY 2939 CTATATATTAATCTCTCAAAACCTACAGCTTTTACTTCACTTCAATATCTTTGAA 2998
DB 3812 CGATGTAGAGATCTTCAGAAATCGCTTCACTCTGTTTTCATTAATCCTGATGG 3753
QY 2999 GAAAGACAAATATCTCAGATCTAATCTAAGACGATCTTACTATTTCCATTAATAT 3058
DB 3752 GTTGTGCAATATTTCCGACATCCACTTTAAGAGGCGATGACCTCCAACTCATAAAT 3693
QY 3059 TTTCCTCTCTGATATCTGATATCTTATTTGCGATTCGATGCTAATCTGTTACT 3118

```

```

DB 3692 TGCCCTTACGCCAATACCGTAAATTAACGATTTGATATCATATGATATCAACAAT 3633
QY 3119 TCATCTATTAATTAACGTCACATGATTTGCTGCTATATATTCGAATTAAGAGTACT 3178
DB 3632 TCATCCAGAAAAGAAAGAAACCAACATTTGGCTTTTCAAAACGCC-----TTGGGTTCT 3579
QY 3179 CTATGACAAATATATTAATTTACTTCTCTACTACCAAAATATATTAACCAAGTAAATTA 3238
DB 3578 TGAGAAAGACACAGTAAAGCACTTTTCAATGACCAAAAGATTCAGATTAATTAATCT 3519
QY 3239 GGTGTAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3298
DB 3518 TTGGGATTCGCTGCTCATTTTGGGCAATTAAGAGCAAAAGCAAGAGGCTTATGACA 3459
QY 3299 TGATTAAGCTTACCACTACTTCTTTCCAAACCTGCGCAACAGTAAATGATATGACA 3358
DB 3458 TGCAATGCGGTAGCAACAGCTTTTCACTGTAACCCGATTCCTCATTAATTAAGACACTA 3399
QY 3359 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3418
DB 3398 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3339
QY 3419 CCGATATTTTCAATATC--CTCAATGCTGATTTCACTGATATTTCTCTACGTAAT 3475
DB 3338 CCAATCATATGAGAAACAGATTAAGATATTTTCAAGATTAATGCGCTGCTTTGCTCT 3279
QY 3476 CTACCAAGCTCTATGCTCTCTTTTCACTAATCTTATCTTCTCTGTAAGAGCTTT 3535
DB 3278 CGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3219
QY 3536 TGTATATATCATTAAGCAACCAATGACAGATTTTACAGAGTGGCAATTTTCAATGAC 3595
DB 3218 GGTATGATTAAGAAAGCACTTTTGAATGATTAATGATGATGATGATGATGATGATGAT 3159
QY 3596 CCACTAATCATTAATTAAGCAACCAATGATGATGATGATGATGATGATGATGATGAT 3655
DB 3158 GAGGCAATTAATTAATTAAGCAACCAATGATGATGATGATGATGATGATGATGATGAT 3099
QY 3656 AATTCATCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3715
DB 3098 AATTCATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3045
QY 3716 TCATTAAGCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3775
DB 3044 CTATTTAGCGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2985
QY 3776 TCATTAATTAATTTTATTAATTTGATGATGATGATGATGATGATGATGATGATGAT 3835
DB 2984 CGTGAATTTGCGCTTCAAGAACCCAGGAAATTAAGCTGATCATCAACCAATCA 2925
QY 3836 TCAGAAATTAATCAATCTTT 3854
DB 2924 TTTCCTTTTGTGATGATGAT 2906

```

RESULT 15
AX599046/c
LOCUS AX599046 8056 bp DNA linear PAT 14-FEB-2003
DEFINITION Sequence 386 from Patent WO02077272.
ACCESSION AX599046
VERSION AX599046.1 GI:28399186
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1
AUTHORS
Berlin, K., Braun, A., Distler, J., Guelzig, D., Howe, A., Mueller, J.,
Olek, A., Piepenbrock, C., Adorjan, P., Grabs, G., Liesch, R., Leu, E.,
Lewin, A., Lipscher, B., Mater, S., Model, F., Mueller, V., Otto, T.,
Piet, C. and Ziebarth, H.
TITLE
Methods and nucleic acids for the analysis of hematopoietic cell
proliferative disorders

QY 2284 TTATGTATATTTATCAATTCACAAATTTCCACTTCAACACAAACTCAAGTCTAT 2343
DB 2702 TATATATATTTTAAATTTTAAATTTATATTTTAA-----AATTTAAAT 2654
QY 2344 TATATCAATTCAGAAATTATATATCTTCAACACACATATGTATCAATTTATATA 2403
DB 2653 TTTAATCAATTTTATATTTTATTTTAAATTTTATATTTTATATTTTATTTTCA 2594
QY 2404 CATCATTCCTTTTAAATAGTACTTCAACATTCATACATAATATATATACATGCA 2463
DB 2593 TTTATTTTATTTTATTTTAAATTTTATTTAAATTTTAAATTTTATTTATTTATA 2534
QY 2464 TGTATTTTCAAGAACATTCACATTCACAGAAATTTTCAATTTAGGTGTATGTGACG 2523
DB 2533 ATTTATTTATTTTATTTTATTTTCAATTAATTTTAAATTT-----ATTTTAAACAAAT 2478
QY 2524 TTCCATTCACAAATTCAGAGTGTGTATCATTTCTCCAAATAGAGTAACTGAGT 2583
DB 2477 AATATATATATTTATTTAAATTTTAAATTTTATTTTAAATTTAAATTTATTT 2418
QY 2584 TTTTAAATCTGTCTTCAATTTCTCAAGCTTTACGTATGTATGATCAGAAATPACTTT 2643
DB 2417 TTTTATTTTATTTTCAAAATTTATTAATTTTAAATTTATTTATTTAAATTTAAATTT 2358
QY 2644 AGCACTTAAATCATTTAATAGGGAATTAGATCTATCTATCTGTGTAAATCTTTGCT 2703
DB 2357 ATATTTTAAATTTAAATTTAAATTT-----TATTTAAATCAAAATTTAATTAATTTTAA 2303
QY 2704 TGTATCATCTCTTACAGATTTTCATATTTAAATTCATTTCTATTAATAGTAAATTTG 2763
DB 2302 AAAATTAATTTAAATTTAAATTTAAATTTTATTTATTTTAAATTTAAATTTAA 2243
QY 2764 TCTTAAAGTACTGCGCATTCATATGACTGCTATTAAGCTTCACTCACTTAATAC 2823
DB 2242 TTTTATTTAAATTTAAATTTTATTTTATTTTATTTTATTTTCAATTTTAAATTTT 2183
QY 2824 ATGAGTACAAACCTATTTTATTTTACAGATCTATTTCAATTAATACCTACAAATTCGG 2883
DB 2182 AATTAATTTATTTTATTTTATTTTATTTTATTTTAAATTTAAATTTAAATTTTAA 2123
QY 2884 TATATCTGACAAATTTCTATTAAGAGGACTCTAAATGGAAGACATTTAATCTATA 2943
DB 2122 AATATATTTAAATTTATTTATTTATTTATTTATTTAAATTTTAAATTTTAAAT 2063
QY 2944 ATATATATCTCACAAACCTACAGCTTTACTCTCACTTCAATTTCTTGAAGAAGA 3003
DB 2062 TTTTAAATTTAAATTTAAATTTAAATTTTATTTTAAATTTATTT----- 2013
QY 3004 CACATTAATTTCTACATCTATTCACAGAACTTACTATTTTCCCTATATATTTTCC 3063
DB 2012 -AAAAATTTATTTAAATTTTAAATTTAAATTTATTTATTTTAAATTTAAATTTATA 1954
QY 3064 CTCCTGATATACTGAGTAATCTTAAATTTGGGTATGCTATGCTAAATCTGTACTTCATC 3123
DB 1953 TTTTATTTATTTATTTATTTATTTATTTATTTTATTTTAAATTTTATTTTAT 1885
QY 3124 TATTAATTAAGTACATATTTGCTGCTATTAATTTCAATTAATAGAGGACTCTATG 3183
DB 1894 AATTAATTAATTTAAATTTAAATTTAAATTTATTTTATTTTATTTTATTTTAA 1835
QY 3184 AGACATATATTTATTTATTTCTGCTACACAAATATATTTACAAAGTAATTTAGTGTG 3243
DB 1834 AATCAATTTTATTTTATTTATTTAAATTTAAATTTAAATTTATTTTATTTTATTT 1775
QY 3244 TAGCATATGATGATATTAATTAATGAGTATCATACCCCTGTATTTTATGTAT 3303
DB 1774 T-----TATTTTAAATTTTATTTTATTTTATTTTAAATTTAAATTTAAATTTATA 1719
QY 3304 TAGCTACAGTACTTCTCTTCCAGACCTGGGAGCCAGTATGAGTATACAGACCTGA 3363
DB 1718 TATTTATTTATTTAAATTTTATTTTATTTTAAATTTAAATTTTATTTATCATTTAAT 1659
QY 3364 TGTATGACTGCTTATTTATTTATTTCTCAATTTAGTATTAACAGGAGGATTAACGAC 3423

DB 1658 TAAATATATTTTATTTTATTTTATTTTCAATTTATTTTCAATTTTAAATTTAAATTT 1599
QY 3424 TATTTCAATATCTCAATATGCTGATTTCACTCATTTCTGTACGTATATCTACAGA 3483
DB 1598 ATTAATATTAATTTAAATTTAAATTTAAATTTAAATTTTAAATTTAAATTTTATTTA 1539
QY 3484 CTCATAGCTCTCTTACAGTACTTATATCTTCTCTGTGAAAGGCTTTCTATATA 3543
DB 1538 AAAATTTTATTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATA 1479
QY 3544 ATCATAGCAATCCATATGACAGACTTACAGAGTGGCAATATTTACATGCCCTAAT 3603
DB 1478 TTTTATATTAATTTATTTTAAATTTAAATTTTATTTTATTTTAAATTTAAATTTATA 1419
QY 3604 CATTAATACAGCAATTAAGATTTACCTTTCTTATG-----CTTTCCAGTACACT 3654
DB 1418 AAAATTTAAATTTAAATTTCAATTTATTTTATTTAAATTTAAATTTTATTTTAAATTT 1359
QY 3655 TATTCATCATATATGATCTCTTACCATATATTCATTAATACATCAGGCTCTT 3714
DB 1358 TTTAAATTTAAATTTATTTTATTTAAATTTATTAATTTATTTTATTTATTTAATTT 1299
QY 3715 TTTATAGGCACTTGATGCGGATTAACCATCATCTGTGTAATTTAGTACATATTTATC 3774
DB 1298 TTTTAAATTTATTTACATTTTATTTATTTATTTATTTATTTATTTATTTATTTCA 1239
QY 3775 ATCACTTAAATATATCTTTTATTTAGATTT-CTGATATCAATTCATCATCAACATTAATA 3833
DB 1238 AAAATTTATTTATTTATTTAAATTTAAATTTAAATTTATTTAAATTTTAAATTTAAATTT 1179
QY 3834 CTTCAGAAATATACATCTTTCTTGGACATTTCAAAATCTCTGTGCAATTTACTCTA 3893
DB 1178 AATATATTAACAAATATATTTATTTTAAATTTTCAAAATTCAAATTTAAATTTAAATTT 1119
QY 3894 ATACTATGAATATCTGCAACACATGACTCTTTTAAAGTACCTCATATATTTACTACC 3953
DB 1118 CAATATATATCAATTTTAAATTTAAATTTAAATTTAAATTTAAATTTTATTTTAAATTT 1059
QY 3954 ACACACATATGATCAACACATATTTAAATTTCAATTTTATTTTACTTTAAATTT- 4012
DB 1058 TTAATATTTAAATTTCAAAATTTAAATTTAAATTTAAATTTTAAATTTTAAATTTT 999
QY 4013 -----AAATATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 4067
DB 998 AAAATTTATTTATTTATTTTATTTTATTTTAAATTTAAATTTTAAATTTTAAATTTTAA 939
QY 4068 CTATATATCAAAATTAATTTAAATTTAAATTTAAATTTAAATTTTAAATTTTAAATTT 4127
DB 938 AAAATTTTATTTCAAAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 879
QY 4128 ACTGATA-CTAATATTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 4186
DB 878 TTTTATTTATTTAAATTTAAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 819
QY 4187 CATGCTATATGTTTCCATTTAAATTTCACTTTAAATTTCACTTTAAATTTTAAATTTTCA 4246
DB 818 CATTTATTAATTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 759
QY 4247 ACGCAGACAGATATGATATCTTTATGAGTAATTTTAAATTTTAAATTTTAAATTTTAA 4306
DB 758 TAAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 699
QY 4307 AATATCCACAAATGATATCAATATGATGATATGATATGATATGATATGATATGATATG 4366
DB 698 AATTTTATTTATTTATTTATTTTATTTTATTTTATTTTAAATTTTAAATTTTAAATTTT 639
QY 4367 ATTTATCATTTGATCTTCAAGTATTAATTTAAATTTAAATTTAAATTTTAAATTTTAA 4426
DB 638 TTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 582
QY 4427 TGAAGTCAATCATCCGTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 4486

[illegible]

Qy	2262	ATCAAAATTAATGAAGAAAATCGTATGATATTAATTAACAATTCACAAATTTCCAACTTCA	2321
Db	2199	ATTTTTTTTTAAAAAAAATTTATTTATTAATTAATTAATTAATTAATTAATTAATTAATTA	2258
Qy	2322	ACAACAAAACCTCAAAGTCTATTATTAACAATTCAGAAATTAATATATCTTCAACCAAC	2381
Db	2259	ATTATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	2317
Qy	2382	ATAGTTTCAAATTCCTTAATATACATCATTCCTTTTAAAGAAATGACTCCACAATTCAT	2441
Db	2318	ATATATTTTAAATGAAAAATATTTTAAATTTTAAATTAATTAATTAATTAATTAATTAAT	2377
Qy	2442	ACATAAATTAATTAATACATGCAATGTTTATTCAAGAACATTACCAATCCAGAAATTTTCA	2501
Db	2378	AATTAATAATA-----AATCTATTCATATTAATTAATTAATTAATTAATTAATTAATTAAT	2432
Qy	2502	ATTACAGGTGAATAGCTGAACGCTCCATCCCAACAAATTCAGAGGTGCTGATACATAC	2561
Db	2433	AACATTAATTAATTAATTAATTAATTTATTTATTAATTAATTAATTAATTAATTAATTAAT	2492
Qy	2562	CTCCAAAACGAGTAATCAAGTCTTTTAATATACGTCTTCAATTCCTCCAGACTTAC	2621
Db	2493	TTTTTTTTTTTTTTTTTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	2552
Qy	2622	GTAAATGTACGAATTAATTAATCTTGAACCTTAACATCATTAATTAAGCGAATTGATCTA	2681
Db	2553	TATGAAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	2612
Qy	2682	TATCTACTGTAATCTTTTGGCTGTAATCATCTCCTTAGAGATTATTAATTAATCC	2741
Db	2613	TATTAATTTTTTTAAAAATTTCTTAAAGTAATTAATTTTACAAAAAATTTATTAATTAATA	2672
Qy	2742	ATCTATTAACATTAACGTAATTCGTCTTAAGTTTACCTCGCATTCATATAGCTGCATTCCTA	2801
Db	2673	AATCATGTTTTTTAAAAAATTAACAAAAATTTTAAATTAATTAATTAATTAATTAATTAATA	2732
Qy	2802	TTAAAGCTTCATCACTTAATACATGAGTAACAAACCTATTTTTCACATGCTATCA	2861
Db	2733	TAAATTTATTAATTT-----TTTCATTTTAAAAAAAATTTTAAAAAAAATTAATTTT	2787
Qy	2862	TAAATATCTTACACAAATTCGGTATATCTGATATTTCTGATATTTCTAAAGCGTATCTTAA	2921
Db	2788	TTTTAAAAAAAACCTATATATCTAATTTAAATTAATTAATTAATTAATTAATTAATTAAT	2847
Qy	2922	TTGAAAGGACATTAAATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	2981
Db	2848	TTAATTAATTAATTAATTAATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	2907
Qy	2982	TTTCAATATCTTTGGAAGAAAGACAAATTAATTCACATGCTATTAACAAGAAATCTTAC	3041
Db	2908	ATATTAATTTTAAAAATTTTAAATTCCTCATTTTATTAATTAATTAATTAATTTTAAATAA	2967
Qy	3042	TATTTTCCCTAATTAATTTTCCCTCGTAATATCTGAGTAATCTTAATTTGGCTATCGT	3101
Db	2968	TAAATTAATTAATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	3026
Qy	3102	ATCGTAATATCTGTACTCTCATCTTAATAATAACGTACACGATTTGCTGCTATTAATTC	3161
Db	3027	ATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTT	3086
Qy	3162	CAATTAATGAGGAGTACTCTATGACAAATTAATTAATTAATTAATTAATTAATTAATTAAT	3221
Db	3087	TTAAATAATTTCTTAATGATATTAATTTTAAATTAATTAATTAATTAATTAATTAATTAAT	3146
Qy	3222	TAACCAAGTAATTAATTAATGCTGTAGCAATGATGAGTACATGATATTAATTAATGAGTATCAT	3281
Db	3147	TTTAAAAAATTAACAAAAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTTT	3206
Qy	3282	ACCCCTGGAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	3341
Db	3207	ATTTTTCATTTTTTTTAAAAAATTTTAAAAAAAATTAATTTTTTTTTAAAAAAAATTAAT	3266
Qy	3342	CAGTAAATGATTAACGACTGATGATGACTGCTTTATTAATCATATCTTCCAAATTAAC	3401

PUBMED 1455227
REFERENCE 8 (bases 14917 to 19517)
AUTHORS Lewis, D.L., Farr, C.L., Farguhar, A.L. and Kaguni, L.S.
TITLE Sequence, organization, and evolution of the A+T region of
JOURNAL Drosophila melanogaster mitochondrial DNA
MEDLINE 94285822
MOL. BIOL. EVOL. 11 (3), 523-538 (1994)
PUBMED 8015445
REFERENCE 9 (bases 1 to 408; 13319 to 19517)
AUTHORS Lewis, D.L., Farr, C.L. and Kaguni, L.S.
TITLE Drosophila melanogaster mitochondrial DNA: completion of the
JOURNAL nucleotide sequence and evolutionary comparisons
MEDLINE 96423163
MOL. BIOL. EVOL. 4 (4), 263-278 (1995)
PUBMED 8825764
REFERENCE 10 (bases 1 to 19517)
AUTHORS Lewis, D.L., Farr, C.L. and Kaguni, L.S.
TITLE Direct Submision
JOURNAL Submitted (03-OCT-1995) Laurie S. Kaguni, Biochemistry Department,
Michigan State University, East Lansing, MI 48824-1319, USA
FEATURES
source
1. 19517
/organism="Drosophila melanogaster"
/organelle="mitochondrion"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/note="derived from new and previously submitted
sequences; sequence is a composite containing sequences
obtained from different Drosophila melanogaster strains"
1. 65
/product="crna-11e"
complement(97. 165)
/product="crna-gln"
171. 239
/product="crna-phe"
240. 1265
/codon_start=1
/transl_table=5
/product="NADH dehydrogenase subunit 2"
/protein_id="AAC47811.1"
/db_xref="gi:1166530"
/translation="MNNSSKILFTITMITITLITVTSNKLGMGLBNLSPFL
LSDNNLSTASLKYFLTVLSTVLSLILMLKNNMNEINSEFTMIISAL
LKGAPPHFEPNMEGLTWNALMTWQKIALPLISYLNKXLLSVILSVII
KLGINGTSLRKMAFBSINHLGMLSLMISSEIMLILFFSFLPTMTFNF
KFLNQLQFSWNSKILKFLFMPLSLGSLPPFLGFLPKMLVIOQLTNCQFMFL
IMMSTLTLFFYLRCISAFMNVFENNIMKNNSINVMIMTFPSIGLFLI
SLFVMP"
1264. 1329
/product="crna-tyr"
complement(1322. 1383)
/product="crna-cys"
complement(1403. 1468)
/product="crna-tyr"
join(1470. 1472,1474. 3009)
/note="mechanism underlying reading frame shift after
first codon uncertain"
/codon_start=1
/transl_table=5
/product="cytochrome c oxidase subunit I"
/protein_id="AAC47812.2"
/db_xref="gi:7412849"
/translation="MSRMTSTNKHQIGTLYTFCGAGAGVGTSLIRAEIGHG
ALIGDQIVITVHATIMLFFMMMPMIGFENGLVPLMGAFFRANMNRN
LIPALSLVSSWENAGATGWTVPPLSGAGASVDLAIFSLHGISLILGA
VNFITVINRSSTGISLDRMLFVWSVITALLLSPLVAGATMLTDRNLNMF
PDPAGGDIILYOHLPFFGGEVYILLPFGVISHISQSGKKEITFSLGKIYAM
LAIGLGIIVAHGFTVGMVDTRAVTSATIMIAVETGKIFSMATLQGLQSLAS
PAIWAIGFVETVAGLTVLANSSVDIILHDTYVVAHFVLSMGAVIMAF
IHWYPLTGLTNMKSHRITIMFGIVLTFEFOHFLGLAGMRASVDPAVTM
IVSTIGSTISLGLIFPFIITWESVSRQYIYIQANSLIENQNPRAHSISELP
LILN"
3012. 3077

CDS
/product="crna-leu"
3083. 3767
/note="TAA stop codon is completed by the addition of 3' A
residue to the mRNA"
/codon_start=1
/transl_except=(pos:3767,aa:TERM)
/transl_table=5
/product="cytochrome c oxidase subunit II"
/protein_id="AAC47813.1"
/db_xref="gi:1166532"
/translation="MSTWANTGIQDSASPLMQLIFFDHALLITVITVGLYEM
LENNYVRPILHGLIEMITIPATILRLPSRLYLDEINPSVTLKSGH
QWMSYSDRNIEFDSYMTPTNEMTDGRILDVDRVLPMSQRIIVTADVI
HMTVPALGVKVDGTGRLNLTNPFIRNPGFLYGCSEICGANSFMEVIESVAVY
FLKWSNNNS"
3768. 3838
/product="crna-lys"
3840. 3906
/product="crna-asn"
3907. 4068
/codon_start=1
/transl_table=5
/product="atpase 8"
/protein_id="AAC47814.1"
/db_xref="gi:1166533"
/translation="MPQMAPISMLLFIIFSIITLIFCSINYSYMPSPKSNELKNI
NINSNMWKX"
4062. 4736
/codon_start=1
/transl_table=5
/product="atpase 6"
/protein_id="AAC47815.1"
/db_xref="gi:1166534"
/translation="MTNLFSVDPDLATPNSINWLSNPLGLMTPSTYIMSPRYNI
KNSILITLHKFKTLDPSCNRSSTPFTISFSLINNNGLFPYFTSRSHTLT
LSLALPMLCEMLVGMNHTQMFAHLPQGPALIMPFCVCIETISYIRGLAVR
LTANMIAGHLTLTLGNSTGSMVLMPLMAQIALVLDSAVAMIOSYFAVAVLTL
VSEVN"
4736. 5524
/codon_start=1
/transl_table=5
/product="cytochrome c oxidase subunit III"
/protein_id="AAC47816.1"
/db_xref="gi:1166535"
/translation="MSTSHNMFHLVDYSPMLTGAIGAMTVSGVYKMFQYDISLF
VGNITITLTVQMRDVSREGTQGLTVAVTIGRMGLIFISEVLPFVSFWAF
PSSISPALIELGASWPMGIISEFNPFOIPLNTAILLAGVTVAAHSHMNNHSC
TQGLPFTVLIGVFTIILQAYEYIEADPTIADSIYGSTPMATGFHIGVILGTFLV
CLIRLNHRSKHNFCEFAAAYMHFVDVYMLFYITTYMGG"
5543. 5607
/product="crna-gly"
5608. 5961
/codon_start=1
/transl_table=5
/product="NADH dehydrogenase subunit 3"
/protein_id="AAC47817.1"
/db_xref="gi:1166536"
/translation="MSITFIALLLITTTVMFLASLKKALIDREKSPFECGPD
PKSSRLPSLRFLLITIIIFLVEIALILPMIIMKYSNIMTWITISITLILLI
GLYHMQMLMSN"
5981. 6045
/product="crna-ala"
6055. 6118
/product="crna-arg"
6119. 6183
/product="crna-asn"
6184. 6251
Query Match 3.9%; Score 184; DB 3; Length 19517;
Best Local Similarity 43.0%; Pred. No. 1,2e-14;
Matches 1844; Conservative 0; Mismatches 2365; Indels 81; Gaps 16;
QY 293 TATTCAGAGATACCTCTTCTTGTTCATCCATATTTCCATATTTTACAAATATCAT 352

Db	15162	TATTAATGATAAATTAACTTTTAAATTTAAAAATTATTTACACTTTTTAA	15221
Oy	353	AAGAGTCCTTATGTGGACATCATCTTAAAATGCATCAACAACAAATTAATGTCTATCA	412
Db	15222	AAATATTTTTTTTTTAAAAAAAATTTTTTTTTTAAAAAAAATTTTTTTTTTAAAAAAAT	15281
Oy	413	ATTATATAAAAAATCCTCATAGTCTTTAGCAGAAGCTACAAATATATAGACACTTA	472
Db	15282	TTTAAAAATTTTAGATTAATTTCTTTTAAATGCTAAAAAAATTTTTTTTTTAAGAT	15341
Oy	473	CTTATTA--ATAATTCCAAATTTATATACCTGCTCATCTATTTAGCATATATCACAT	530
Db	15342	TTTAAAACTTTTTTTTACAAATTTTTTAAAAAATATATATATTTAAATTTTAAAAAAT	15401
Oy	531	CTGACATGATTAACAATAGTAACACATTTCTTAACACTAGATATATACAAATTAATTA	590
Db	15402	TTTTTTTTTAAAAAAAATGA AAAATTAATATATAAAAATTTTTTTTTTCAAAAAATGAAA	15461
Oy	591	TATATTTTTTACGTGTTCTTGATAAATTTAGTAATAAAATGAAAATCAGATTTAG	650
Db	15462	TTTATATCATTTAAAAAAAATTTATTAATTTTATGAAATATATAAAAAGTATTAAT	15521
Oy	651	GTCAGAGATCTTTGGACAGCAATAGCAATTCACCTGCAGCAGCATGGTATATCGATTA	710
Db	15522	TTATTTAAAAATCAATATATATATATATTAATTAATTTGATATTAATTAATTAACGA	15581
Oy	711	ACTATGGGGGCGATCATAGAAATATACACTATTAACCTTCCGAAAATTTTA	770
Db	15582	TAAATATATATATTAATTTATTTTATATCCTAAATCTGAATATATATATATATA	15641
Oy	771	AATATTTAACCATCATCATCTACAGACAACATATATAGCAACGCAATATAGCGAAG	830
Db	15642	TATATATATATATATA-----TACATATATATATATATATATATCAATTAATTAATAAT	15696
Oy	831	TATATATCGAACAAATACATAGTATATCTTAACTATCTCCACAGCAATTAACGCCA	890
Db	15697	AAATTTATCCCTATTCATTAATTTATTAATTTAAATTTAAACTTBAABAAGTATTTTT	15756
Oy	891	TATGTACAAATTAACAACAACAACAGCATATGTGTA AAAATATCTCCAATATTTAATTTGTA	950
Db	15757	TTAAAAAAAATATATTTATTTAAATTTATCTTAATTAATTAATTTTATTAATTAATTTT	15816
Oy	951	GTAAGGTATCGAAATACATCACTCAAAATTTCCGATGAAATAGAGAGAAATTTTAC	1010
Db	15817	ATATATTAATTTATTTAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	15876
Oy	1011	AATATTAATCCAAATTTTATATCTCTGTGTCCAAAGTTTCTTAAGAAATTCGAGAACATC	1070
Db	15877	ATAATTTTTTAAAAAATTCATTTATATTTATATATATATATATATATATATATTAATTT	15936
Oy	1071	TTCCCTGTAGTATAGTACTGCTGTGGATATATTAAGAACTGTGTGATCATGTATAGAA	1130
Db	15937	TCATTTAAATTAATTAATTAATTAATTAATTTATTTTAATCACTAAATCTGAATATAT	15996
Oy	1131	CAATTAAGTATGATGTTCTTAAAAATATATACCATCAAGATATTAAGGTACAGATTTG	1190
Db	15997	TAAATTAAT	16051
Oy	1191	GAGCTGATTAAGAAACATATATGCAATGCAATGTGATTAATCGGTGAAAAAATTTAG	1250
Db	16052	-TTCAATTAATTTATGTATATTTAAACTTAAAAATATTTTTTTTTTAAAAAAAATGCA	16110
Oy	1251	GTAATTAATGCTGTTGCTACTGTTATTAACATAAGCATGAATGAAATTA AAACATAATTA	1310
Db	16111	TTTATTTAAATTAATCTTAATTAACATTTTATTAATTAATTTATTTATTAATTAATTAAT	16170
Oy	1311	TAGCAAAAAATCATCAATGATCTTCATACATTAATTTGTCATCATGCTTGGAGATC	1370
Db	16171	TTTAAATTAATTAATTAATTAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTA	16230
Oy	1371	TAAATTTAATATGTACACAGAACATTCACGCATATGCTTTGGACTAGAAATAGGA	1430

[illegible]

QY 2502 ATTTCAGTGTGATGCTGACCTTCATTCACAAATTCAGACGTTGTGTATCATTTAC 2561
Db 17349 AACTAT 17408
QY 2562 CTCGAAAGAGATGATGAGTGTATTAATAGCTGCTTCAATCTCAGAGCTTTAC 2621
Db 17409 TTTTGT 17468
QY 2622 GATATGATCAGAAATGATCTTACGCTTAAATCATTAATGAGCAATGTAGTACTA 2681
Db 17469 TATGAT 17528
QY 2682 TATCTATGAT 2741
Db 17529 TAT 17588
QY 2742 ATTCTAT 2801
Db 17589 AATCATGTTTAT 17648
QY 2802 TTTAGCTTCATCCTTAAATGATGATGACAAATCTATTTTTCAGATGCTATTTCA 2861
Db 17649 TAT 17703
QY 2862 TAAAT 2921
Db 17704 TTTTAT 17763
QY 2922 TTGGAAGGACATTTAT 2981
Db 17764 TAT 17823
QY 2982 TTTCATATCTTGGAGAGACACATATATCTGACATCTATCTACAGAAATCTTAC 3041
Db 17824 AT 17883
QY 3042 TATTTTCCCTAT 3101
Db 17884 TAT 17942
QY 3102 ATCGAT 3161
Db 17943 AT 18002
QY 3162 CAATATGAGAGATCTATGAGACAAATATATATATATATATATATATATATATAT 3221
Db 18003 TTTAAAT 18062
QY 3222 TAAACAAGTAT 3281
Db 18063 TTTAAAT 18122
QY 3282 ACCCTGTGATTTTAT 3341
Db 18123 ATTTTCAATTTTAT 18182
QY 3342 CAGTAT 3401
Db 18183 AT 18242
QY 3402 GTATATACAGGAGATGAGATATATATATATATATATATATATATATATATATAT 3461
Db 18243 ATATCTAT 18295
QY 3462 TTTCTATGAT 3521
Db 18296 TAT 18355
QY 3522 CTGTAATAGGCTTTCTAT 3581
Db 18356 TGATTTAT 18415

QY 3582 CAAT 3641
Db 18416 AT 18475
QY 3642 TTTCCAGTAT 3697
Db 18476 TCTTAT 18535
QY 3698 ACAACATGAGGCTTTTCTATATATATATATATATATATATATATATATATATATAT 3757
Db 18536 TAAACAAAAATTTTAT 18595
QY 3758 TTAGATGAT 3817
Db 18596 TTTTAT 18655
QY 3818 TCATCAACAAT 3877
Db 18656 TTTAT 18715
QY 3878 GCGAT 3935
Db 18716 AATTTAT 18775
QY 3936 ACACCTCAT 3995
Db 18776 CTGAT 18835
QY 3996 TTTTAT 4055
Db 18836 TAT 18895
QY 4056 AGCAT 4115
Db 18896 AATTTAT 18944
QY 4116 TCATCTTGTAACTAT 4175
Db 18945 AATGAT 19004
QY 4176 CCGTGAACAGATGCTAT 4235
Db 19005 AAAAAATTTTAT 19064
QY 4236 TTCAAT 4295
Db 19065 TAAAAAATTTTAT 19124
QY 4296 CTGAT 4355
Db 19125 AATTAAT 19184
QY 4356 CCAACAAGAT 4415
Db 19185 AAT 19244
QY 4416 TCTCAT 4468
Db 19245 ATTTAT 19304
QY 4469 ATGTCACAAAT 4528
Db 19305 ATTTAT 19364
QY 4529 ATATCTTGAATCTTAT 4588
Db 19365 AAT 19394

RESULT 18
AC093899/c 172816 bp DNA linear PRI 12-JUN-2002
LOCUS AC093899 Homo sapiens BAC clone Rpl1-724016 from 2, complete sequence.
DEFINITION

ACCESSION AC093899 AC068884
 VERSION AC093899.3 GI:18497265
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS Sulston, D.E. and Waterston, R.
 TITLE Toward a complete human genome sequence
 JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
 MEDLINE 99063792
 PUBMED 9847074

REFERENCE
 AUTHORS Pearman, C., Haakenson, W. and Boyer, E.
 TITLE The sequence of Homo sapiens BAC clone RP11-724016
 JOURNAL Unpublished (2001)
 REFERENCE 3 (bases 1 to 172816)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (10-SEP-2001) Genome Sequencing Center, Washington University School of Medicine, 444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE 4 (bases 1 to 172816)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (05-FEB-2002) Genome Sequencing Center, Washington University School of Medicine, 444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE 5 (bases 1 to 172816)
 AUTHORS Waterston, R.
 TITLE Direct Submission
 JOURNAL Submitted (12-JUN-2002) Department of Genetics, Washington University, 444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On Feb 5, 2002 this sequence version replaced gi:15625013.

REFERENCE
 AUTHORS Center: Washington University Genome Sequencing Center
 TITLE Center code: WUGSC
 JOURNAL Web site: http://genome.wustl.edu/gsc
 REFERENCE Contact: sapiens@wustl.wustl.edu
 JOURNAL Summary Statistics
 REFERENCE Center project name: H_NH0724016
 JOURNAL Drafting Center: WIBR

COMMENT
 NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
 Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see http://genome.wustl.edu/gsc

SOURCE INFORMATION:
 The RP11-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moon, P.Y., Zhao, B., Frengen, E., Taten, M., Catane, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong and coworkers at http://www.chori.org
 VECTOR: pBACE3.6

NEIGHBORING SEQUENCE INFORMATION:
 Actual start of this clone is at base position 1 of RP11-724016,
 actual end is at base position 172816 of RP11-724016.

Unresolved simple sequence repeat from base position 22700 to 25900.

FEATURES
 The sequence of AC068884 has been incorporated into AC093899.

source	location/Qualifiers
repeat_region	1..172816
repeat_region	19..1140
repeat_region	1141..1437
repeat_region	1438..2657
repeat_region	2658..2729
repeat_region	2730..3320
repeat_region	3321..3642
repeat_region	3643..4184
repeat_region	4185..4996
repeat_region	5011..5263
repeat_region	5285..5907
repeat_region	6895..7078
repeat_region	7103..7421
repeat_region	7807..7952
repeat_region	7981..8289
repeat_region	8360..8551
repeat_region	8622..8718
repeat_region	8784..8825
repeat_region	8840..9134
repeat_region	9136..9174
repeat_region	9861..9881
repeat_region	9882..10166
repeat_region	16502..16556
repeat_region	16886..17165
repeat_region	17179..17482
repeat_region	18025..18119
repeat_region	18177..18471
repeat_region	19312..19441
repeat_region	19731..20040

```

repeat_region      /rpt_family="MaLR"
20403 . .20472
repeat_region      /rpt_family="L2"
21286 . .21344
repeat_region      /rpt_family="CA)n"
22724 . .23977
repeat_region      /rpt_family="(TA)n"
24005 . .24171
repeat_region      /rpt_family="(TA)n"
24196 . .25007
repeat_region      /rpt_family="(TA)n"
25030 . .25205
repeat_region      /rpt_family="(TA)n"
25223 . .25398
repeat_region      /rpt_family="(TA)n"
25399 . .25455
repeat_region      /rpt_family="AT_rich"
25456 . .25919
repeat_region      /rpt_family="(TA)n"
25920 . .26206
repeat_region      /rpt_family="Alu"
26893 . .26917
repeat_region      /rpt_family="AT_rich"
27194 . .27497
repeat_region      /rpt_family="L1"
27498 . .27814
repeat_region      /rpt_family="MER2_type"
27815 . .28151
repeat_region      /rpt_family="L1"
28152 . .28278
repeat_region      /rpt_family="Alu"
28279 . .28485
repeat_region      /rpt_family="L1"
28795 . .28829
repeat_region      /rpt_family="AT_rich"
28866 . .29100
repeat_region      /rpt_family="L2"
29237 . .29349
repeat_region      /rpt_family="L1"
29350 . .29484
repeat_region      /rpt_family="Alu"
29485 . .29589
repeat_region      /rpt_family="L1"
29611 . .29874
repeat_region      /rpt_family="L2"
29875 . .30284
repeat_region      /rpt_family="L1"
30320 . .30522
repeat_region      /rpt_family="L1"
30523 . .30847
repeat_region      /rpt_family="Alu"
30848 . .30882

```

[illegible]

QY	538	GATTACAAATGTGAACACTTCTTCAACACTGATATAATATACATATAATATATATAT	597
DB	25594	TATATAAAATATATAAAATGTATATATATAATATATATATATATATAATATAATAT	25533
QY	598	TTGACGTGTTCTTGA--TATATAATAGTAAAAAATGAAAATCAGATATTTAGTGC	654
DB	25594	ATATATATATATATATATATGTATATATAATATATATATATATATATAATATATATA	25475
QY	655	AGATTCATTTGGCAGACAAATGCAATTTGCACTGTGAGACATGATATCTAGTTACCT	714
DB	25474	ATATATATTTTATATATATATAATATATATAATTTATATATATAATATATATAATAC	25415
QY	715	ATGGGACGTGATCATAAGAAATATTCACATATAAACCTTACCGAAAAAATTTAAATA	774
DB	25414	TATATATATTTATATATAATTTTATATAATATATATATCTTTATATATATAATATAT	25355
QY	775	TTTACCCACATATCATCTACAGACAACTATTTGCAACGACGATATATAGAGGATAT	834
DB	25354	TATATATATATATATATATTAACATATATATATATATATATATATATATATATAT	25295
QY	835	ATCTGACAAATATCATATGATATATATCTTAACTATCTTACACAAATTAAGCAATATG	894
DB	25294	ATATATATTAACATATAATATATATATATATATATATATATATATATATATATAT	25235
QY	895	TACCAATATCAACACAAACAGCATATATGTATTAATCTCCAAATTTATTTTATAT	954
DB	25234	ATATAATATATTTAAACATATAATATATATATATATATATATATATATATATAT	25175
QY	955	AGGATATGAAATTAATCATCTCAATATTTCCAGTGAATAGCAGAGAAATTTTACAATA	1014
DB	25174	ATTAATATATAAATATATATATATTAACGTAAATATATTTATATGATATAATATTTAA	25115
QY	1015	TATTCATTTTATACTCTCTGCTCAAGTTTGGTATAGAAATTTGCAACATCTTCC	1074
DB	25114	AAATATATATTTATATATATATATATATATATTAACGTAAATATATATATATATATA	25055
QY	1075	TTTGATATATGACTTGGCTGTGATATATAAGACTTGTGAATCATTTGATGAACAAT	1134
DB	25054	TATTAACATATAATATATATATTTATGATATAATATATTAACGTAAATATATATTTAT	24995
QY	1135	AAGTATGATGTTCTAAAAATATATATACATCAAGATTTATAGGTGCAGATTTGAGC	1194
DB	24994	TATATAATATATTTAAACGTAAATATATATTTATCTATTAACATATATCAATATATA	24935
QY	1195	TGCAATTAAGACATATATTGCAATTTGCATGTGGAATATTCCTGTGAAAAATTTAGCTA	1254
DB	24934	TATTAATATATATAAATATATATATATCAAAATGTAATATATATATATATATATAT	24875
QY	1255	TAAATGCTGTGACGTGATATACTAAAGCATGAAATGAATTAACACTATATATAGC	1314
DB	24874	CAATGTAAATATATATATATATATATATATATATCAAAATGTAATATATTAATATA	24815
QY	1315	AAAAAATCATTCATATGATCTTCATATCATTAATTTGGTCACTATGCTTTGAGATCTAT	1374
DB	24814	TAAAAATATATCAAAATATATAATATATATCAAAATATATAATATATATTT--ATATGTA	24758
QY	1375	ATTACATGTATCAACAGAACTTTCAAGCAATATGCTTTTGGACATAGAAATAGAAAAAG	1434
DB	24757	TATATATATAATATATATATATATATATGATATATAAACATATATATATATATATA	24698
QY	1435	TAGAAATTTAAATACATTAATATATATCAAACTTAAAGCTTGTGAAGAACCATATCTGT	1494
DB	24697	TTTATATATATAACATATTTAAATATCA---TATATATATATATATATAATATATATA	24643
QY	1495	AAAAACATGATATCATTTGCAAAAAAACTTATATGACATCAACCAATTTGCATATCTAT	1554
DB	24642	ATATCAATATATATTTAT	24583
QY	1555	TTTCAATTTTATTCATGAGATATATATCTGATATAAGCATATCAATCATATATATCTTA	1614
DB	24582	ATTTTAAATACAAATGTATATAAATATATATATATTTATATATAATATATAACATATA	24523
QY	1615	GTATATATATCTTTGATATATTTCAACACATGATATAAACGTGCAAAATATATGATA	1674

D	24522	TATTTATATTAACGTATATAAATATATAATATATTTATATATTAAGTATATAAATATATAATATA	24463
OY	1675	CAAGCTTCACCAAGCATATATAAAGAGCGTCAAAATCCCATCGTTAAATCCCACTAAGTA	1734
D	24462	TTTATATTAACGTATATAAATATAATATATTTATATTAACGTATATAAATATAATATAT	24403
OY	1735	TGTAAACATATTACATAACAACTTAACTTAACCTTAAATTTCTAGCTATTATATATAA	1794
D	24402	TATATTAACGTATATAAATATATAATATATTTATATTAACGTATATAAATATAATATATTA	24343
OY	1795	CAAAATTCCTATCTCTCATTCATTAACAAATGTTATCAAGATTATATAAATAGTCAAAAT	1854
D	24342	TATTAACGTATATAAATATATAATATATTTATATTAACGTATATAAATATAATATAT	24286
OY	1855	AAACCTTTACTTATTTATAAATATATGTTTGACACCTTTTCACTACATCATGTATAT	1914
D	24285	ATATTAACGTATATAAATATAAATATATTTATATTAACGTATATAAATATAATATTTATAT	24222
OY	1915	AATGAACATATTTTACATATAAATACTAATATAAAGAACCAATACATATATTTAAATAC	1974
D	24225	AACGTATATAAATATAAATATATTTATATTAACGTATATAAATATAATATATATATAT	24166
OY	1975	TAAACACCTTTTACTACAGCAATATATTAACATAGATGCTCGACATATATAATCTATCAT	2034
D	24165	AAAATATATAAATATAAATTTTATTAATATATAAATAATATATAATATAATATAATTA	24106
OY	2035	CCTTGACATATTTCAAAATATAAATATGATTAACGTGTTTTTTATGCCCATTAGTTCTTA	2094
D	24105	TATATTAATATATATAAATATAAATATATAAATATAAATATATATATATATAATATAAAT	24046
OY	2095	ATATAAATTCCTAAATTAACCAAAAAAGTTAAATAACATTCCTCGACACATATGAAGAAG	2154
D	24045	ATATAATATATTAATATATAATATATTAATATAAATATATTAATACAT--AAATAT	23985
OY	2155	ATACATACACGTCAAATATAATTAACAAAATTTCTAATATTAACACCTACTTAAACATA	2214
D	23988	ATTATAAATATAAATATCTATATATAAATATATAAATATAAATATATTAAGTATATATA	23928
OY	2215	GGAACCAATGATATATCATACCTAAGTAAGATATCCCTACAGATATCAATAATATAA	2274
D	23928	ATATATTAATATAAATATAATATTTATATATTAATAATATAATATATTAATATAA	23865
OY	2275	GAAAAATCGTATATATTTATTAACAATTAACAAATTTCCACCTCGACACACAAACCTCA	2334
D	23868	TATATAATATATGTTTATATATATAATATAATATATATTAAGTATATATTCACATATA	23809
OY	2335	AAGTCTATATATAACAATTCAGAAATTAATATATCTTCACCAACATATGTATCAAT	2394
D	23808	TATATATATTAATATAAATATAATATTTATATATATAATATAATATAATATATTTTATA	23749
OY	2395	TCTATATACAT-CATATCCTTTTAAAAAGATGACCTTCACATTCATACATAAATAT	2453
D	23748	TATATATAAATATAATTTATATATTTATAATATAATATAATATATATGAATATATAATAT	23689
OY	2454	ATACATGCAATGTTTATTCGAAACATTAACAATTCAGAAATTTTCAATTATACGTGTA	2513
D	23688	ATATATGATATATATAAATATAAATATAATATATATATATAAATAAATAATATATAT	23629
OY	2514	ATGCTGAACGTTCATCCCAACAAATTCAGCAGTGTGATCATTAACCTCCAAAACGAG	2573
D	23628	ATATTAATATAATATAATATATTAATATTTAAATATATTTAATATAATATAATATAA	23565
OY	2574	ATTAACGAGTTTTTAAATATCTGCTTTCAAAATCTTCAGAGCTTTAGCTAATGCTACAG	2633
D	23568	TATTAATATATTTAATAATATAATATAATATATAAATAAATAATATAAATAATAT	23509
OY	2634	AAATTAACCTTAGACATTAACATCATTAATAGCAAAATTAAGATCATATATCTACGTGTA	2693
D	23508	ATTTAATATATATATATAATATAATATAATATATTTAATAAATAATATAATATAAATA	23449
OY	2694	AACTTTTGCTGTAATCATCTCCTTAGAGAGATTTCAATATAAATCCATTCATATACAT	2753

Db	23448	ATATATGTTATATATATATAT-----ATATTTATATATATATATATATATATATAA	23397
Qy	2754	TACGTAATAGTCCTTAGTACCGGCGCATTCACATGACGATGGTATTAAGCTTCAT	2813
Db	23396	TATATGTTTATATATATATATATTAATATATAATATATATATATATATATATATGTTTAT	23337
Qy	2814	CACCTTAATACATGATGACAAACCTATTTTTTTCAGATGCTATTCATATAAATACCTAC	2873
Db	23336	ATATATATATATATTAATATATAATATATGTTTATATATATATATATATAATATAAATAT	23277
Qy	2874	ACATATCCGGATATCTGTACATATTTCTACTTAAGAAGGCTCCTTAATGGAGACAT	2933
Db	23276	ATGTTTATATATATATATATTTAAATATAATATATATATTAATATATATATATTAAT	23217
Qy	2934	TTA-----ATCTATATATTAATCCTCACAAAACCTACCAAGCTTTTACTCACTTCAAT	2988
Db	23216	ATTAATATATATATTTAAATATATATTTAAATATAATATATTTATATATAATATTAAT	23157
Qy	2989	ATCTTTGAAGAAGACACATATATTTCTACATCTATCTACACAGAAATCTTACATTTTC	3048
Db	23156	ATTAATATATATATATATATAATATTTTAAATATATAATATATTTTACATATGATATATAT	23097
Qy	3049	CCTATATATTTT 3061	
Db	23096	TAAATATATAATAT 23084	

RESULT 19	LOCUS	DEFINITION	AE001398/c
AE001398	14867 bp	DNA	linear
AE001398	3D7 chromosome 2	section 35 of 73	of the
AE001398	complete sequence.		
AE001398.1	GI:3845197		

KEYWORDS	SOURCE
Plasmodium falciparum	

ORGANISM
Plasmodium falciparum 3D7
Eukaryota, Alveolata; Apicomplexa; Haemosporida; Plasmodium.
1 (bases 1 to 14867)

AUTHORS
Gardner, M.J., Tetteh, H., Carucci, D.J., Cummings, L.M., Arzvin, L., Koonin, E.V., Shellen, S., Mason, T., Yu, K., Fujii, C., Pederson, J., Shen, X., Jing, J., Aston, C., Lai, Z., Schwartz, D.C., Perera, M., Salzberg, S., Zhou, L., Sutton, G.G., Clayton, R., White, C., Sitt, H.O., Fraser, C.M., Adams, M.D., Venter, J.C. and Hoffman, S.L.

TITLE
Chromosome 2 sequence of the human malaria parasite *Plasmodium falciparum*

JOURNAL Science 282 (5391), 1126-1132 (1998)
MEDLINE 99021743
ISSN 0036-8075

PUMED
 9804551
 2 (bases 1 to 14867)
 Gardner, M.J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R.W.,
 AUTHORS
 REFERENCE

Carlton, J. M., Pain, A., Nelson, K. E., Bowman, S., Paulsen, I. T., James, K., Eisen, J. A., Rutherford, K., Salzberg, S. L., Craig, A.,

Angiolini, S., Perlea, M., Allen, J., Selengut, J., Haft, D.,
Kyes, S., Chan, M.-S., Nene, V., Shalim, S.O., Sun, B., Peterson, J.,

Mañner, M. W., Valdivia, A. E., Martin, D. M. A., Fairlie, A. H., Fraunholz, M. J., Roos, D. S., Ralph, S. A., McEadden, G. I.,

Cummings, L.M., Sudarmanian, G.M., Mungall, C., Venter, J.C., Carnucci, D.J., Hoffman, S.L., Newbold, C., Davis, R.W., Fraser, C.M. and Barvelli, D.

TITLE
Journal
Genome sequence of the human malaria parasite *Plasmodium falciparum*
Nature 410 408-511 (2002)
DATE/ED.

BOOKS	199-211	1992
REFERENCE	3	(bases 1 to 14867)
AUTHORS	Gardner M. T.	

TITLE Submitted (02-OCT-2002) The Institute for Genomic Research, 9712
JOURNAL

Medical Center Dr, Rockville, MD 20850, USA
location/Qualifiers

```
source
1. .14867
/organism="Plasmodium falciparum 3D7"
```

```
/mol_type="genomic DNA"
/isolate="3D7"
```

/db_xref="taxon:36329"

/note="AT_rich"
/rpt_type=tandem
complement(5437..5484)
/note="AT_rich"
/rpt_type=tandem
complement(5554..5577)
/note="AT_rich"
/rpt_type=tandem
complement(6302..6350)
/note="AT_rich"

Query Match 3.8%; Score 182.2; DB 3; Length 14867;
Best Local Similarity 45.4%; Pred. No. 2.2e-14;
Matches 1284; Conservative 0; Mismatches 1478; Indels 65; Gaps 15;

271 TCTCTTACATTAAGAAACAGCTGATTCAGAGATACCTCTCTCTGTCATCCATATT 330
9811 TTTTAAATATGTTCAATATATATTTTACAAAGTGAATTAATCATGTAACCTATTA 9752

331 TTCCAAATTTTACATATCATAGAGTTCTTTATTCGATTCATATTAAATCATCA 390
9751 TTCCGAAATTTTATAGAAAACATATAAGTTAGACATAAATATTATGTAATA 9692

391 AACAAATATATAGTACTATCTATTTAAATTAATAAAATCTCATAGCTTTAGCAAGAACT 450
9691 TTACAGATATATATCATTTGTAATAATTAATAATTTATTTAATAATCATATAT 9632

451 ACAATATATAGACACATTAATTTATATATTTTCAATTTATTTACTACTATCA 510
9631 CATTTTTCATTCAGAGTGTAAAT--TTTATATTTTATTTTAAATTTATAT 9575

511 TATTAGCATATATCATCTGACATGATTAACATAGTAACACATTTCTTAACACTAG 570
9574 TATTTATATATTAATTTATGTTATTAATTAATAAATTAATATATATATATAT 9515

571 ATATATATATATATATATATATATATTTTGAATGCTGTTCTTGATATAATTTAGTAA 630
9514 ATATATATATATATATATATATATTTTAAATTTT--AGTATATATTTTAAATTTATAT 9456

631 AATGAAATCAGTATTTTGTGACAGATCATTTGACAGCAATGCAATGCAATGCACTGC 690
9455 AAGGATATATTTTATTAATTAATTAATGATGTTAAATTAATTAATTAATTA 9396

691 AGCAGATGATATATGTTAACTTATGAGGAGCTGATCATAGAAATATTAACATATA 750
9395 AAAAAATATTTTAAATATATTTTAAATTAATTAATTAATTAATTAATTAATTA 9336

751 CACTTACGAAATATTTAAATTTTACCCATAT--TCATTCACAGACATATTA 807
9335 TATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 9276

808 TGCAACAGCATATAGAGAGATTTATCTGACACATATACATGATTTATCTTAACAT 867
9275 AATTAATTTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 9216

868 TCCATACACATATAGCAGCATATGTACACAAATTAACACAAACGCAATGTGTA 927
9215 AATTAATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 9156

928 AATATCTCCATATATTTGTAGTAAGATTCGAATTTACATCATCAATTTCCAG 987
9155 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 9096

988 TGAATAGCAGAAATTTTACATATATATCAATTTTATCTCTGCTGCAAGTT 1047
9095 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 9036

1048 TGCATAAGAAATTTGCAACATCTCTGTAGTATAGTACTGCTGATTAATTAAGA 1107
9035 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 8976

1108 ACTTGATATCATGATTAAGAAACATTAAGTATGATGTTCTTAATTAATTAATTA 1167

8975 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 8921
1168 AGATATATATGATGTCAGATGAGCTGATTAAGAAACATTAATGCAATGCAATG 1227
8920 ATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 8861

1228 AATATGCTGAGAAATTTAGTAATAATGCTGTTGCTACTGTTAATTAAGGCA 1287
8860 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 8802

1288 GAATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1347
8801 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 8742

1348 TGTGTCATCATGCTTGTGAGATCTAAATTAATTAATTAATTAATTAATTA 1407
8741 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 8685

1408 GCGTTTGTGATGAATTAAGAAAGAGTAATTAATTAATTAATTAATTAATTA 1467
8684 ACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 8625

1468 AAGCTTGTGAAGAACAGTACTGTAACCACTGATATCATTAAGCAAAATCTTA 1527
8624 AATATATAT-----AATTAAGATTAAGCAATTTTATTTTAAATTAATTA 8572

1528 TGTGAATCTACCAATTTGCAATCTATTTTCAATTTTATTAATGAGATTAATCACTGA 1587
8571 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 8512

1588 TAAAGCATATCAAAATATTTATCTAGTCATCATCTTTGATTAATTAATTAATTA 1647
8511 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 8452

1648 TAAATTAAGCTGCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1707
8451 TTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 8392

1708 TACCAGATGTTACATCACTAGATTAATTAATTAATTAATTAATTAATTAATTA 1767
8391 TATATATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 8333

1768 TTAATTAATTTCTAGTAAATTAATTAATTAATTAATTAATTAATTAATTA 1825
8332 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 8273

1826 TGTGTCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1885
8272 ATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 8213

1886 TGCACTTTTCTACATCATGATTAATTAATTAATTAATTAATTAATTAATTA 1945
8212 AATTAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 8153

1946 AAAAAACCAATACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2005
8152 ATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 8093

2006 AGATGCTGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2065
8092 -----TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 8041

2066 AACTGTTTTTTATGCCCATTAATTAATTAATTAATTAATTAATTAATTAATTA 2125
8040 ATTAATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 7981

2126 ATACATATCTGCAACATTAATTAATTAATTAATTAATTAATTAATTAATTA 2185
7980 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 7927

2186 TTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2245
7926 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 7867

QY 2246 TATCCCTACAGATGATCAATATATAAGAAAAATCTATATATTTATACATTTAC 2305
 DB 7866 TTTTAT 7807
 QY 2306 AAAATTTCCATCTCAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 2365
 DB 7806 TAT 7747
 QY 2366 TATATCTCCACACATATGTTATCAATCTATATATATATATATATATATAT 2425
 DB 7746 TAT 7693
 QY 2426 TCACTTCATATCTATATATATATATATATATATATATATATATATATATAT 2485
 DB 7692 TAAATTAAT 7633
 QY 2486 AATCCAGAAATTTCAATTTACGATGATATGCTGACAGCTTCACAAATTCAGCA 2545
 DB 7632 ATGTAAT 7586
 QY 2546 GTTCGTGATACATATCTCCAAACAGATATACGATTTTAAATATCTGCTTCAAT 2605
 DB 7585 TTTCTTTTAT 7526
 QY 2606 TCTTCACAGCTTTACGATATGATACAGAAATATACCTTAAACATCATTAATA 2665
 DB 7525 ATAAATATACAGCTGATATATATATATATATATATATATATATATATATAT 7466
 QY 2666 GCGAT 2725
 DB 7465 AAT 7406
 QY 2726 TTTCAAT 2785
 DB 7405 ATTAAT 7346
 QY 2786 TATGATGAT 2845
 DB 7345 TAT 7286
 QY 2846 TTACAGATGAT 2905
 DB 7285 TTA-----TTAGATATATATATATATATATATATATATATATATATATAT 7231
 QY 2906 AAAGACGATCTC-TAATGGAGAGATATATATATATATATATATATATATATAT 2964
 DB 7230 AATTAAT 7171
 QY 2965 ACCAGCTTTTACTGATCTTCAATATCTTTGAGAGAGACAAATATATCTCACATCTAT 3024
 DB 7170 TATGAT 7111
 QY 3025 ACTACAGAGATCTTACTATTTTCCCTATATATATTTTCCCTCGTAAATATCTGAGTAA 3084
 DB 7110 TCTAAT 7051
 QY 3085 TCTTAAT 3091
 DB 7050 TTTAAT 7044

RESULT 20
 AE015944 300205 bp DNA linear BCT 06-FEB-2003
 LOCUS Clostridium tetani E88, section 9 of 10 of the complete genome.
 DEFINITION AE015944 AE015927
 ACCESSION AE015944.1 GI:28204325
 VERSION
 KEYWORDS
 SOURCE Clostridium tetani E88
 ORGANISM Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 Clostridium.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCES
 AUTHORS
 TITLE
 JOURNAL
 FEATURES
 source
 1 (bases 1 to 300205)
 Bruggemann, H., Baumer, S., Fricke, W.F., Wierzer, A., Liesegang, H.,
 Decker, I., Herzberg, C., Martinez-Arias, R., Merkl, R., Henne, A., and
 Gottschalk, G.
 The genome sequence of Clostridium tetani, the causative agent of
 tetanus disease
 Proc. Natl. Acad. Sci. U.S.A. 100 (3), 1316-1321 (2003)
 2 (bases 1 to 300205)
 Bruggemann, H., Baumer, S., Fricke, W.F., Wierzer, A., Liesegang, H.,
 Decker, I., Herzberg, C., Martinez-Arias, R., Merkl, R., Henne, A., and
 Gottschalk, G.
 Direct Submission
 Submitted (31-OCT-2002) Goettingen Genomics Laboratory, Institute
 of Microbiology and Genetics, Georg-August University,
 Grisebachstr. 8, Goettingen 37077, Germany
 Location/Qualifiers
 1. 300205
 /organism="Clostridium tetani E88"
 /mol_type="genomic DNA"
 /strain="Massachusetts"
 /sub_strain="E88"
 /db_xref="taxon:212717"
 /complement (162. 683)
 /locus_tag="CTC02285"
 /complement (162. 683)
 /locus_tag="CTC02285"
 /note="tip-3"
 /codon_start=1
 /transl_table=11
 /product="bacterial protein translation initiation factor
 3"
 /protein_id="AA036764.1"
 /db_xref="GI:28204326"
 /translation="NNIIKSGFLMNEEIRBEKRVYRIADGEOGVIPSEALKRAEEK
 ELDIVIAPTGKPPVCKIMNYGKPIEDKKDKKAKKKOKYINVKIETLSATIEHDI
 GIKANNARKFLAEEDNKVYTVAFRGREHSHSVNGKILKTLFSKVEDCVYKEKARL
 ESKNMTVLAERK"
 /complement (865. 2784)
 /locus_tag="CTC02286"
 /complement (865. 2784)
 /locus_tag="CTC02286"
 /EC_number="6.1.1.3"
 /codon_start=1
 /transl_table=11
 /product="threonyl-tRNA synthetase"
 /protein_id="AA036765.1"
 /db_xref="GI:28204327"
 /translation="MINITLKGKVIYEKGVKVDIYMKISPALYKXAVGAKINGEI
 AELMEIKEDSELILTFDDERGRKTVHTSHLQAVKRLYPKALIGPAIDNGF
 YDPPIDITFTPEMKEIKEMAKIVKNSLEIERELPREBAIKVDSAPKVEYKRI
 EDLPGEVISTFYKQDFVLDLQAGPHLPSTGKIKAIKLSVNGATYRGDEKMKLORIT
 ENPWEKHKTRAGVDEIRTPVILNELIWEHRSQHWMDPKENMTFTKIDNSNPAIKPNVCP
 GSILVYKSHLSHYKEFPWRLGLVHRSBGLHGMARVACFQDDAHAFMTLEOI
 KDEILNVILKIDSFYKFGFEYFVLSTRPEDSMGSDDEVAATNGLXNLAEGAELEY
 KINEDGAFYEGKIDFHLKDCIGRTWOCGTICLDQWPERFPLTVGGDGEKRGVVM
 HRVFGSTIERFGILIEHFAGAPTWLAPOVKWTITDSOKVKNVINDIKKGIKIR
 VEPDRNEKIGKIRBAQIQKVPYITILGDKEVSNKVAVSRKSGDLAGLSKEPVA
 KLYEIDNRIVENSK"
 /complement (3117. 4016)
 /locus_tag="CTC02287"
 /complement (3117. 4016)
 /locus_tag="CTC02287"
 /codon_start=1
 /transl_table=11
 /product="conserved protein"
 /protein_id="AA036766.1"
 /db_xref="GI:28204328"
 /translation="MKYMAIKIIVEDDPKEIIBGINALKYFPELEIGVCSIDNINH
 FWRIVYPEKTLKSKSFIHLIANIIVNIYIDKLEKMTPEIESYFNSDKDIKE
 IBEKRDILTNNKMKIDSSNIYVKKKQEIIMDEIIECKENNEINIEGFLFRKYLE

gene
CDS
TRNA
tRNA
tRNA
CDS
gene
CDS
gene
CDS
gene
CDS

KELKLIIEIKVEEYAEKYDEDFILKLKFVNIQSRIIDEVNIIYYEGGANIIDKRSKI
IIDITTELFDPDIISNNWSLEINDDLITLSGLIVGCEKIIIHVCENCLKEIINIINW
FLKKVAQNSGCETGRINKQYKR"
4093..4812
/locus_tag="CTC022288"
4093..4812
/locus_tag="CTC02288"
/codon_start=1
/transl_table=11
/product="conserved protein"
/protein_id="AAO36766.1"
/db_xref="GI:28204329"
/translation="MLGVTKYDIPVDGRISTEEFYNI.MSLSENIICKPIESLLYAI
SHPDLHVLDEKNLILHKDIDKEFIKKLNUNILNKSQSSJNSNPNGIGLVANVA
LSNVPIHNKITYDPLVENVKDKNLINVGKTCSAVVNDNAITSDKIAQAKRE
VSDIVLLTPPGDIIPLGLDYGFIGGCCGILRDLMAFPFDLNTFNYGNEIKEPILYN
VPPIYLKKKGLIDRGSILLNE"
complement(5057..5131)
/product="tRNA-Cys"
/note="codon recognized: UGC"
complement(5138..5212)
/product="tRNA-Gly"
/note="codon recognized: GGC"
complement(5218..5293)
/product="tRNA-Phe"
/note="codon recognized: UUC"
complement(5403..5491)
/locus_tag="CTC02290"
complement(5403..6491)
/locus_tag="CTC02290"
EC_number="2.4.2.21"
/codon_start=1
/transl_table=11
/product="nucleotide-nucleoside-dimethylbenzimidazole
phosphoribosyltransferase"
/protein_id="AAO36768.1"
/db_xref="GI:28204330"
/translation="WTIMKLBETLRIRIKPADEVIKKANSRNDISKPIGSKLKE
ETRAVISATIGTCNKELINKRMVILWCSSNGICEBSVSCPELTALLAENVKEITGI
EVALSHRVAGDCVVDIGA.KSDVKOSRI.LANKVAVGTKNMAEPMTREEALAEITSI
BLVDFFHVGEADLPFGTGAGIGNTSAIVSVLSQIDSKIYKSGGLTEBGLINKR
RAYKAKIALINKNEKEDVIDIAKIGSPFAICIGCFGLGAARVPIYDIGIISASLL
CAKKAANYKDELILLSHAEBGIEYTVGALGKRCPLMEMRLDEBGCCPLPFMIIES
ALCLNNMATTAAEASIVDSSEFVDIRER"
complement(6976..7875)
/locus_tag="CTC02291"
complement(6976..7875)
/locus_tag="CTC02291"
complement(6976..7875)
/locus_tag="CTC02291"
/codon_start=1
/transl_table=11
/product="hsp33 (redox regulated chaperone)"
/protein_id="AAO36769.1"
/db_xref="GI:28204331"
/translation="EKGKMDAPLKIKAKDGQVRITGAILTELWNKGIEVHKSPRG
AAALSRMTLAGLMSMLKSEKDITITRIDGGEEKGLVTAPREGVAKGITGNPLVH
LPLANGGLVDVGA.IKGNKITVILKDWPKOPYIGQVYSGEIGDLAYFTVSQT
PSAVLNGVLYVDRKDISIKASGFEILOWMGADELADPLYTLBEIPSTIELYSKMSI
BELIFPEIGHMDKILBIGIVPEYTCDSREREKIDRALISIGYDKIETEYDEGKTEELVC
OFENEKRYFVDEKHICELIRIMNN"
complement(7879..8619)
/locus_tag="CTC02292"
complement(7879..8619)
/locus_tag="CTC02292"
EC_number="2.1.1.-"
/codon_start=1
/transl_table=11
/product="methyltransferase"
/protein_id="AAO36770.1"
/db_xref="GI:28204332"
/translation="MECKYDFAOIYDNLIREDVVYHSMSNFIMDIYTHVWKDNYLD
IACCGKTEIISKFKTTNAVDSLDAEAEMKLYXNGRGAFIQODISKILEKGE
PNLITGCCDINSIVLTISEVESYFLRVYTNHDKGIFIFDINSYTKITEVANNITYY
DSEEFYVMENTLAKDIOVENTLTFIKKANSYFERFDEVHERARYKRWLEKLIKONN

	gene	FKLIDKNGEYFAVEEKSRIVYYLVKK"
		9133..10131
		/locus_tag="CTC02293"
CDS		9133..10131
		/locus_tag="CTC02293"
		/EC_number="1.2.1.11"
		/codon_start=1
		/transl_table=11
		/product="aspartate-semialdehyde dehydrogenase"
		/protein_id="AA036771.1"
		/db_xref="GI:28204333"
gene		/translation="MCKRYNALVCGTGWGQKFLILERRKPPVNVNLIFFASRAG KKIQLCEKEFTTBELKDNIKNKKIIDIALNSAGGITSTETPAIFVRIGATVDSSSA RMDSPLVPEVNPDVDKLTKGIINPNCSTIQAVALKPIDEGIRIVISTQA VSAGAAGVNDLDKGNGVAPKFEPVALGNALPHIDSFDNGYTEEMKAMVETRKI FNHEGKITATVRIPIVFNHSESIIVELEKEFEDVELVKNVGIIITLEDDEVLY YPMAITAGSDQDVFVIRIRDPSFINNGNLMLVVADNRKGAALNTIQIAELIKNK" 10158..11042
CDS		/locus_tag="CTC02294"
		10158..11042
		/locus_tag="CTC02294"
		/EC_number="4.2.1.52"
		/codon_start=1
		/transl_table=11
		/product="dihydrodipicolinate synthase"
		/protein_id="AA036772.1"
		/db_xref="GI:28204334"
gene		/translation="MSIFPTSGVAITYTPCEGNNPKULBELTEHIENSTDAIVGL TTBASTMTBAZIKETIKFTVDQGRIPVIYACGSNNTRKAIELSKAKSITDVGLL LTIPIFNKTIOKGVIEHFRAINDVDVPFIILVNPBTGTNLIFKLVICDSNV ALEAEGDLSQIITEMKALLKRLIDIVSGDDQIIPISIGIGIVSLVINIMEKHVD MCBLVYKGTGKALBIQLNLTSLTNLSLETNPIPVKTAMINGDIVGEIRUPLCNMD ESNINLKELSHNALIK" 11057..11824
CDS		/locus_tag="CTC02295"
		11057..11824
		/locus_tag="CTC02295"
		/EC_number="1.3.1.26"
		/codon_start=1
		/transl_table=11
		/product="dihydrodipicolinate reductase"
		/protein_id="AA036773.1"
		/db_xref="GI:28204335"
		/translation="MEVYLIMRVLMIVGNGXMGKRITTHCSKPNDIEIYAGVDKSTS NDPFPVENIHSSVECDVLDPSRBSLSLISYTEKCLPVLCTTGYSKEELNLI
Query Match	3.8%; Score 182; DB 1; Length 300205;	
Best Local Similarity	49.3%; Pred. No. 8.1e-15;	
Matches 606; Conservative	0; Mismatches 600; Indels 24; Gaps 4;	
Oy	2498 TTCAATTACGGGTAATGCTGAAGCTTCATCCACAATAATCAGACGTTGCTATAC	2557
Db	45599 TTTATTTTAACTTTAAAGTTCTTCAACTCATCTTAGAATTCAGCTGTTTTCTCC	45655
Oy	2558 TTAACCTCAAACAGATTAACGAGTTTAAAATGCTGCTCTTCAAAATCTTCACAGCT	2617
Db	45659 TTTCACATGGAAAGTTTGGATTAATTCATATACCCTTTTTCATAGTGTTTAAAGCT	45718
Oy	2618 TTACGTAATGTAACAGAAATTAATTAGAACCTTAAACATCATTAATAGGCATATAGAT	2677
Db	45719 TCCTCCTATATATAGAAATTAATCATCTTTTCTCTATCATATATTCGTTTGAAAT	45778
Oy	2678 ACTAATCTACTGTGTAATCTTTTGCTGTAAATCATCTCTTACAGAGATTTCATATTAA	2737
Db	45779 A---TCCTCCAGAGTAATCTT---CTTTTATTAATTAATGCTCTGTAGACATATT	45833
Oy	2738 ATCCATTCTAATACATTACGTAATGCTCTTAAGTACCTGCCATTCATAGCTCATTT	2797
Db	45833 GCCCTTTCTAATATATTTTGAGCTCTCAACATTTCCATGATATTTATTTCTCAAG	45892
Oy	2798 GCATTTAAAGCTTCAATCACTTAATATATGACATGACAAACATATTTTAAAGATGCTA	2857
Db	45893 CTATCTAAAGCATCCCAGATATCTCTTTTATATCTTTTCTAATTTTAAATTAATATACCTT	45955

QY 2858 TTCTAAATACCTACCAATTCGGGTATCTGTACAAATTTCTACTAAAGCGTACT 2917
 DB 45953 TTTATTAAGATTCTATTAATTAAGCAATATCGCTTTTCTCGCCTTAATGGTGGCAAT 46012
 QY 2918 CTAATTTGGAAGACATTTAATCTAATTAATTAATTCCTACAAAACCTACAGCTTTTACT 2977
 DB 46013 TCTATATTTATTAACATTTATTTCTAATAAACAATCTGTCTAATAATTTTCTTAT 46072
 QY 2978 TCACCTTCAATATCTTTGGAAGAAACACAAATTTCTCATCTATCTAACAAGAAC 3037
 DB 46073 TCTTCTAATTAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 46132
 QY 3038 TTACATATTTCCATTAATTTTCCCTCCGTAATATCTGAAATCTTAATTCGTA 3097
 DB 46133 CTAAACCTATAGCTTTATCTCTTATCTCTTAAACCTTTAAAGTTAGTTGAAA 46192
 QY 3098 TCGATACGTAATCTGTACTTCAATTAATTAATTAATTAATTAATTAATTAATTA 3157
 DB 46193 TTTATGTCATATCCCATTTTCATCTAAACAAATCTCCCTTTGAGATGTTCAAT 46252
 QY 3158 ATTCCAAATAGAGAGTACTCTATGAGCAATATATATTTACTTTCCGACTACCAAT 3217
 DB 46253 ATACCTTTTATCTCTACGCTCCAG-----TAAAGCTCCTTTTATGACCAAT 46306
 QY 3218 ATATTAACCAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3277
 DB 46307 AATTCACCTTTCTAACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 46366
 QY 3278 TCATACCCCTTGATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3337
 DB 46367 TCTTTTCAATTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 46426
 QY 3338 GAACCAAGTATAGTATACGACTCGATGATGATGATGATGATGATGATGATGATGAT 3397
 DB 46427 TCTCAGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 46486
 QY 3398 TTACGTATTAACAGGAGCTTACCGCATTTTCAAT-----AATCTCAATGCT 3445
 DB 46487 TTGTTCATTTCTTTGAAATTTCCATTAATTAATTAATTAATTAATTAATTAATTA 46546
 QY 3446 GATTCACATCATTAATTTCTCTAGTATCTACAGCTTATAGCTCTCTTACACT 3505
 DB 46547 CTGCTTAATATTTTATTTGAAATTTAGATTTATATTTATTTATTTGCTTAGAAT 46606
 QY 3506 AACTTAATCTCTCTCTGTAAGAGCTTTCTTAATTAATTAATTAATTAATTAATTA 3565
 DB 46607 AAACCTAATCTCTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 46666
 QY 3566 GACCTTAACAGAGGAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3625
 DB 46667 GCTTTATACAGAGATTTCTATTTGCTTCAATAGCTGCTATTAATTAATTAATTT 46726
 QY 3626 TACCTTTCTTAAGCTTTTCCAGTACACTTAATCAATCAATCAATCAATCAATCAAT 3685
 DB 46727 TTTATTTCTTTATTTCTTTAGATTTTCCAAACCAATCTCTGACCAATCTTAATCC 46786
 QY 3686 AATTCATTAATTAACAGATGAGCTTTT 3715
 DB 46787 AATTAATTAATTAATTAATTTTCTTTAT 46816

RESULT 21
 AC134919/c 145992 bp DNA linear PRI 07-AUG-2003
 LOCUS DEFINITION Homo sapiens BAC clone Rpl3-485B10 from 4, complete sequence.
 AC134919
 AC134919.3 GI:33386922
 VERSION AC134919.3
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE

1 (bases 1 to 145992)
Sulston, J.E. and Wilson, R.

Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)

99063792

2 (bases 1 to 145992)
9847074

McNellan, M., Haglund, K., Du, H., and Cotton, M.
The sequence of Homo sapiens BAC clone Rpl3-485B10

Unpublished (2001)

3 (bases 1 to 145992)
Waterston, R.H.

Direct Submission
Submitted (02-OCT-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA

4 (bases 1 to 145992)
Waterston, R.H.

Direct Submission
Submitted (20-NOV-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA

5 (bases 1 to 145992)
Wilson, R.K.

Direct Submission
Submitted (01-AUG-2003) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA

6 (bases 1 to 145992)
Wilson, R.K.

Direct Submission
Submitted (07-AUG-2003) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Aug 1, 2003 this sequence version replaced gi:2510186.

----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC

Web site: <http://genome.wustl.edu>
Contact: sapiens@watson.wustl.edu

----- Summary Statistics
Center project name: H_FH0485B10

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=

30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. Wes Warren,
Department of Genetics, Washington University, St. Louis MO. For
additional information about the map position of this sequence, see
<http://genome.wustl.edu>

SOURCE INFORMATION:
The RPL3-13 Human Female BAC Library was constructed using improved
cloning techniques developed by Kazuo Osoegawa. The library was
generated by Baochun Zhao in our laboratory. Construction was
funded by a grant from the National Human Genome Research Institute
(NHGRI, NIH) (#R01HG01165-03). The library was generated according
to the new NHGRI/DOE 'Guidance on Human Subjects in Large-Scale DNA
Sequencing'.

Female blood was obtained via a double-blind selection protocol.
Female blood DNA was isolated from one randomly chosen donor (out
of 10 female donors) and partially digested with a combination of

Db	130042	TAAATATATATGTAATATGTAATGTTATTAATATATATACATATACAGTAATATGAAATGTTAT	129988
Qy	2896	AATATCTACATGAAGAGCGTCTCTATATGGAAGACATTTATCTATATATTAATCTC	2955
Db	129982	AATATATATATATATATATGTAATGTAATGTAATGTTATTAATATATATATATATATATGTAATG	129922
Qy	2956	ACAAACCTACACAGCTTTTACTCTACCTTCCATATCTTTGGAGAAGACACATATATCT	3015
Db	129922	TTATG	129864
Qy	3016	CACATCTATATACACAGCATCTTACTATTTTCCCATATATTTTCCCTCCGTAATTA	3075
Db	129862	AATATTTTAAAT	129803
Qy	3076	CTTGAGTAATCTTAATTTGGGTATCGATCGTAATCTGTACTTCATCTATTA-ATAACG	3134
Db	129802	AATCAAAATTTAT	129744
Qy	3135	TACATGATATTTGCTGCTCATATATCCAAATATGAGAGCTCTATGACAAATATAT	3194
Db	129742	TATATATGTAAT	129688
Qy	3195	TATTAATCTTCCCTACATCAACAAATATATATTAACAGTAATTAATGCTGTAGCATAGTG	3254
Db	129682	AATATATATATTTAT	129622
Qy	3255	AGTACATAGATATATATAT 3271	
Db	129622	AATATATATATATATATAT 129606	
RESULT 22			
AC127383/c		122168 bp	DNA linear PRI 15-OCT-2002
LOCUS			
DEFINITION	Homo sapiens BAC clone RP11-1388P24 from 2, complete sequence.		
ACCESSION	AC127383		
VERSION	AC127383.4	GT:22758593	
KEYWORDS	HTG.		
ORGANISM	Homo sapiens (human)		
SOURCE	Homo sapiens		
REFERENCE	Bukacynski, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 122168)		
REFERENCE	2 (bases 1 to 122168)		
REFERENCE	3 (bases 1 to 122168)		
REFERENCE	4 (bases 1 to 122168)		
REFERENCE	5 (bases 1 to 122168)		
REFERENCE	6 (bases 1 to 122168)		
REFERENCE	7 (bases 1 to 122168)		
REFERENCE	8 (bases 1 to 122168)		
REFERENCE	9 (bases 1 to 122168)		
REFERENCE	10 (bases 1 to 122168)		
REFERENCE	11 (bases 1 to 122168)		
REFERENCE	12 (bases 1 to 122168)		
REFERENCE	13 (bases 1 to 122168)		
REFERENCE	14 (bases 1 to 122168)		
REFERENCE	15 (bases 1 to 122168)		
REFERENCE	16 (bases 1 to 122168)		
REFERENCE	17 (bases 1 to 122168)		
REFERENCE	18 (bases 1 to 122168)		
REFERENCE	19 (bases 1 to 122168)		
REFERENCE	20 (bases 1 to 122168)		
REFERENCE	21 (bases 1 to 122168)		
REFERENCE	22 (bases 1 to 122168)		
REFERENCE	23 (bases 1 to 122168)		
REFERENCE	24 (bases 1 to 122168)		
REFERENCE	25 (bases 1 to 122168)		
REFERENCE	26 (bases 1 to 122168)		
REFERENCE	27 (bases 1 to 122168)		
REFERENCE	28 (bases 1 to 122168)		
REFERENCE	29 (bases 1 to 122168)		
REFERENCE	30 (bases 1 to 122168)		
REFERENCE	31 (bases 1 to 122168)		
REFERENCE	32 (bases 1 to 122168)		
REFERENCE	33 (bases 1 to 122168)		
REFERENCE	34 (bases 1 to 122168)		
REFERENCE	35 (bases 1 to 122168)		
REFERENCE	36 (bases 1 to 122168)		
REFERENCE	37 (bases 1 to 122168)		
REFERENCE	38 (bases 1 to 122168)		
REFERENCE	39 (bases 1 to 122168)		
REFERENCE	40 (bases 1 to 122168)		
REFERENCE	41 (bases 1 to 122168)		
REFERENCE	42 (bases 1 to 122168)		
REFERENCE	43 (bases 1 to 122168)		
REFERENCE	44 (bases 1 to 122168)		
REFERENCE	45 (bases 1 to 122168)		
REFERENCE	46 (bases 1 to 122168)		
REFERENCE	47 (bases 1 to 122168)		
REFERENCE	48 (bases 1 to 122168)		
REFERENCE	49 (bases 1 to 122168)		
REFERENCE	50 (bases 1 to 122168)		
REFERENCE	51 (bases 1 to 122168)		
REFERENCE	52 (bases 1 to 122168)		
REFERENCE	53 (bases 1 to 122168)		
REFERENCE	54 (bases 1 to 122168)		
REFERENCE	55 (bases 1 to 122168)		
REFERENCE	56 (bases 1 to 122168)		
REFERENCE	57 (bases 1 to 122168)		
REFERENCE	58 (bases 1 to 122168)		
REFERENCE	59 (bases 1 to 122168)		
REFERENCE	60 (bases 1 to 122168)		
REFERENCE	61 (bases 1 to 122168)		
REFERENCE	62 (bases 1 to 122168)		
REFERENCE	63 (bases 1 to 122168)		
REFERENCE	64 (bases 1 to 122168)		
REFERENCE	65 (bases 1 to 122168)		
REFERENCE	66 (bases 1 to 122168)		
REFERENCE	67 (bases 1 to 122168)		
REFERENCE	68 (bases 1 to 122168)		
REFERENCE	69 (bases 1 to 122168)		
REFERENCE	70 (bases 1 to 122168)		

JOURNAL

COMMENT

Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Sep 7, 2002 this sequence version replaced gi:22203890.
Genome Center
Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu/gsc>
 Contact: aspiens@wustl.edu
 ----- Summary Statistics -----
 Center project name: H_NH1388724

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCL1 human BAC library was made from the blood of one male donor, as described by Oosagawa, K., Moon, P. Y., Zhao, B., Frengen, E., Tateo, M., Catenean, J. J., and de Jong, P. J. (1998). An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pteret de Jong and coworkers at <http://www.chori.org>

PIRector: p3acc3.6

NEIGHBORING SEQUENCE INFORMATION:

This sequence is not the entire insert of the clone. This clone is overlapped by AC015969 and AC130709.

The sequence from base 48271 to base 48342 was derived from one plasmid subclone.

The sequence from base 69077 to base 69140 was derived from one plasmid subclone.

SOURCES

```

organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone="RP11-1386F24"
/clone_lib="RPC1-11"
5897..6028
/rpc_family="MER1_type"
13148..13310
/rpc_family="1u"
13766..13983
/rpc_family="MR"
14058..14187
/rpc_family="MER1_type"
14764..14842
/rpc_family="MER2_type"
14847..14889
/rpc_family="(T)n"
14850..14526
/rpc_family="MER2_type"

```


[illegible]

[illegible]

FEATURES	source
JOURNAL	Journal of Bacteriology
TITLE	Genomic analysis of the oral bacterium Fusobacterium nucleatum strain ATCC 25586
REFERENCE	1. Bacteriol. 184 (7), 2005-2018 (2002)
REFERENCE	2. (bases 1 to 12277)
REFERENCE	3. (bases 1 to 12277)
REFERENCE	4. (bases 1 to 12277)
REFERENCE	5. (bases 1 to 12277)
REFERENCE	6. (bases 1 to 12277)
REFERENCE	7. (bases 1 to 12277)
REFERENCE	8. (bases 1 to 12277)
REFERENCE	9. (bases 1 to 12277)
REFERENCE	10. (bases 1 to 12277)
REFERENCE	11. (bases 1 to 12277)
REFERENCE	12. (bases 1 to 12277)
REFERENCE	13. (bases 1 to 12277)
REFERENCE	14. (bases 1 to 12277)
REFERENCE	15. (bases 1 to 12277)
REFERENCE	16. (bases 1 to 12277)
REFERENCE	17. (bases 1 to 12277)
REFERENCE	18. (bases 1 to 12277)
REFERENCE	19. (bases 1 to 12277)
REFERENCE	20. (bases 1 to 12277)
REFERENCE	21. (bases 1 to 12277)
REFERENCE	22. (bases 1 to 12277)
REFERENCE	23. (bases 1 to 12277)
REFERENCE	24. (bases 1 to 12277)
REFERENCE	25. (bases 1 to 12277)
REFERENCE	26. (bases 1 to 12277)
REFERENCE	27. (bases 1 to 12277)
REFERENCE	28. (bases 1 to 12277)
REFERENCE	29. (bases 1 to 12277)
REFERENCE	30. (bases 1 to 12277)
REFERENCE	31. (bases 1 to 12277)
REFERENCE	32. (bases 1 to 12277)
REFERENCE	33. (bases 1 to 12277)
REFERENCE	34. (bases 1 to 12277)
REFERENCE	35. (bases 1 to 12277)
REFERENCE	36. (bases 1 to 12277)
REFERENCE	37. (bases 1 to 12277)
REFERENCE	38. (bases 1 to 12277)
REFERENCE	39. (bases 1 to 12277)
REFERENCE	40. (bases 1 to 12277)
REFERENCE	41. (bases 1 to 12277)
REFERENCE	42. (bases 1 to 12277)
REFERENCE	43. (bases 1 to 12277)
REFERENCE	44. (bases 1 to 12277)
REFERENCE	45. (bases 1 to 12277)
REFERENCE	46. (bases 1 to 12277)
REFERENCE	47. (bases 1 to 12277)
REFERENCE	48. (bases 1 to 12277)
REFERENCE	49. (bases 1 to 12277)
REFERENCE	50. (bases 1 to 12277)
REFERENCE	51. (bases 1 to 12277)
REFERENCE	52. (bases 1 to 12277)
REFERENCE	53. (bases 1 to 12277)
REFERENCE	54. (bases 1 to 12277)
REFERENCE	55. (bases 1 to 12277)
REFERENCE	56. (bases 1 to 12277)
REFERENCE	57. (bases 1 to 12277)
REFERENCE	58. (bases 1 to 12277)
REFERENCE	59. (bases 1 to 12277)
REFERENCE	60. (bases 1 to 12277)
REFERENCE	61. (bases 1 to 12277)
REFERENCE	62. (bases 1 to 12277)
REFERENCE	63. (bases 1 to 12277)
REFERENCE	64. (bases 1 to 12277)
REFERENCE	65. (bases 1 to 12277)
REFERENCE	66. (bases 1 to 12277)
REFERENCE	67. (bases 1 to 12277)
REFERENCE	68. (bases 1 to 12277)
REFERENCE	69. (bases 1 to 12277)
REFERENCE	70. (bases 1 to 12277)
REFERENCE	71. (bases 1 to 12277)
REFERENCE	72. (bases 1 to 12277)
REFERENCE	73. (bases 1 to 12277)
REFERENCE	74. (bases 1 to 12277)
REFERENCE	75. (bases 1 to 12277)
REFERENCE	76. (bases 1 to 12277)
REFERENCE	77. (bases 1 to 12277)
REFERENCE	78. (bases 1 to 12277)
REFERENCE	79. (bases 1 to 12277)
REFERENCE	80. (bases 1 to 12277)
REFERENCE	81. (bases 1 to 12277)
REFERENCE	82. (bases 1 to 12277)
REFERENCE	83. (bases 1 to 12277)
REFERENCE	84. (bases 1 to 12277)
REFERENCE	85. (bases 1 to 12277)
REFERENCE	86. (bases 1 to 12277)
REFERENCE	87. (bases 1 to 12277)
REFERENCE	88. (bases 1 to 12277)
REFERENCE	89. (bases 1 to 12277)
REFERENCE	90. (bases 1 to 12277)
REFERENCE	91. (bases 1 to 12277)
REFERENCE	92. (bases 1 to 12277)
REFERENCE	93. (bases 1 to 12277)
REFERENCE	94. (bases 1 to 12277)
REFERENCE	95. (bases 1 to 12277)
REFERENCE	96. (bases 1 to 12277)
REFERENCE	97. (bases 1 to 12277)
REFERENCE	98. (bases 1 to 12277)
REFERENCE	99. (bases 1 to 12277)
REFERENCE	100. (bases 1 to 12277)
REFERENCE	101. (bases 1 to 12277)
REFERENCE	102. (bases 1 to 12277)
REFERENCE	103. (bases 1 to 12277)
REFERENCE	104. (bases 1 to 12277)
REFERENCE	105. (bases 1 to 12277)
REFERENCE	106. (bases 1 to 12277)
REFERENCE	107. (bases 1 to 12277)
REFERENCE	108. (bases 1 to 12277)
REFERENCE	109. (bases 1 to 12277)
REFERENCE	110. (bases 1 to 12277)
REFERENCE	111. (bases 1 to 12277)
REFERENCE	112. (bases 1 to 12277)
REFERENCE	113. (bases 1 to 12277)
REFERENCE	114. (bases 1 to 12277)
REFERENCE	115. (bases 1 to 12277)
REFERENCE	116. (bases 1 to 12277)
REFERENCE	117. (bases 1 to 1

```

gene
  /sub_species="nucleatum"
  /db_xref="ATCC:25586"
  /db_xref="taxon:190304"
  116..859
  /gene="FN0892"
  116..859
  /gene="FN0892"
  /EC_number="3.1.3.3"
  /codon_start=1
  /transl_table=11
  /product="Phosphoserine phosphatase"
  /protein_id="AAL95088.1"
  /db_xref="GI:19714455"
  /transl_table=11
  /translation="MKIKMIAAFDIDGTIYRNALLIEHFKLVKVELPDDIOYPLKV
  EAVYIMPTRGDYDDYLLDTOLYYVAKLPLPKYNDPISNOYILLKGRVYTYRE
  MTEWKKHGHVFTSGSPFLVSRMAKKGVDFGSLVPEIDETVTSFKGLTKPMW
  DSAHKQEAIEHFIKKYNIDLSKYAYGDTNGDFMSLVSNPAINSKELITRVND
  ENLSKTIQIIEERKVIYKLSNDELIEF"
  894..1316
  /gene="FN0893"
  894..1316
  /gene="FN0893"
  /codon_start=1
  /transl_table=11
  /product="Hypothetical protein"
  /protein_id="AAL95089.1"
  /db_xref="GI:19714456"
  /transl_table=11
  /translation="MKIKNEVVALQIIYYLTLRDKDIISSNEISAENIPLPLCIR
  IIKLEKAGVYIKRGAAGVYLRDRKRLFRPIETIIDDYVLQPCIDISSTICSTR
  GADCSIRALKKIQDDLDPDKINFIDLVENNASLQI"
  1347..1883
  /gene="FN0894"
  1347..1883
  /gene="FN0894"
  /codon_start=1
  /transl_table=11
  /product="unknown"
  /protein_id="AAL95090.1"
  /db_xref="GI:19714457"
  /transl_table=11
  /translation="MDIQENKYYADDMGKGVNKLIAQNGAILLYLILPLFLVK
  NGILGSDGTFRLIVCTIFLFLNPLSQKTLFLKTKASIHNDQSYPTLVQFGN
  IFMOGKPSVELDSKIYKLYKSYVLMTFNSNGIMIKYDSPTKCNFDPFEPFIKE
  NCKKAKIIVKRSYIFGL"
  complement(1929..2393)
  /gene="FN0895"
  complement(1929..2393)
  /gene="FN0895"
  /codon_start=1
  /transl_table=11
  /product="unknown"
  /protein_id="AAL95091.1"
  /db_xref="GI:19714458"
  /transl_table=11
  /translation="MDTKLKSITDSFPKGIIDKTIENTYKNHDEKSYSGCRIQEGYD
  DYLLKIVFKRKIEYKDDPFWSDTPDEIIFQELNEVKRDPVEEIVPEIKFEEEN
  DILDKTYKEFFIENNERKKELSKMENYIKVEYIPDEKAMDEENKLLI"
  complement(2417..2800)
  /gene="FN0896"
  complement(2417..2800)
  /gene="FN0896"
  /codon_start=1
  /transl_table=11
  /product="Hypothetical protein"
  /protein_id="AAL95092.1"
  /db_xref="GI:19714459"
  /transl_table=11
  /translation="MKKEKXVVEVLRGGYKVSSEVDLIAQLIPSPFGSPFTIIPKOV
  FLAYNKEFFVIAATLKGDDPKIKIHYSGLDEVNQMKGILLNGULTIHTSDGIRE
  NYRAMKPAFGSLAANNFKALQKFG"
  2903..3571
  /gene="FN0897"
  2903..3571
  /gene="FN0897"
  /codon_start=1
  /transl_table=11

```

```

gene
  /product="Hypothetical cytosolic protein"
  /protein_id="AAL95093.1"
  /db_xref="GI:19714460"
  /transl_table=11
  /translation="MRELKDPFKNKKIDFKKLEFGFPLIDNSYYHTSLNKPQFMS
  VKINDNSIFTEIIDETNEPYILILEMKQSGVSEKYYKAYNEVILIOKECFEDFI
  EKAYNTKRIIDVYKKGDELFFLMKSPKNAVIRKRSNMGVYVLTISRKIGLDS
  DKIEIVNLHNSSEIRKLIIDCKKIFPAPVANKTKWCSICDGYVELEIYKLIIDISY
  ELAK"
  3646..5517
  /gene="FN0898"
  3646..5517
  /gene="FN0898"
  /EC_number="4.1.99.-"
  /note="spore photoproduct"
  /codon_start=1
  /transl_table=11
  /product="Hypothetical protein"
  /protein_id="AAL95094.1"
  /db_xref="GI:19714461"
  /transl_table=11
  /translation="MLYIVTALYIEAKPLISLNLKKNQTYTKQVSENIKLITISG
  TKIKSATATLYLSENKIKNDYITINGFIASSNNNSQIDYIYISKIQAASPTTP
  YPEMYKXNPLBGLITPDKITBKTIENVEYIDMEAYGFOTASIFFKDKTILKIV
  SDILKEVEDRLIDFKODNLFNKS YKTYVPLKFINI PNNKXNNFNNSODLTKVY
  LENKLSDTMYTEFFNIIKYLKIKYGNIDILKXENIEVNSKVGKFIIEEIKESKL
  NNKVEIERKSPFNKNSNLFNNRFSHIYVEKTIINNKTLEISFKDVKYIIEIDNKE
  VSSNNODFHLOKLGOKLILASNKPNTIYEGAVVCESENNFYTSIINCVDCEY
  CYLGIVSGNIVIFVDIEKVEEVEELYSKLTLYLCVSYDTDLAIENIGSFKKW
  YHFIEDKVKIKLELTETKSGNIDKPLNKILPMLNFIIFATSEBNALAKNKYTFASFR
  VKAIEIDPKMKVRIICDPLIYSNFKANSQIMEYFNSIEDKXVIDISIGVRIS
  KEYLKNRQNNSEILYFPPECIDGYTTSYDCKTSYWINFTKQFLEKYNINKLYT"
  5527..6243
  /gene="FN0899"
  5527..6243
  /gene="FN0899"
  /EC_number="1.1.1.100"
  /codon_start=1
  /transl_table=11
  /product="3-oxoacyl-[acyl-carrier protein] reductase"
  /protein_id="AAL95095.1"
  /db_xref="GI:19714462"
  /transl_table=11
  /translation="MKIALVGTSGTSGIGYEISKRLKNYTVYIGRNFINDENIFK
  EYENFIPTVCDLSKDLKELTSLKIKKQDLVNSAGIYFGGHEBNVSKIKMTIT
  VNLOAPVITISQVFLRTKENGKILITINSVANNKSPASVYSATKAGLSQSKLFE
  EYKNDVYITITPMTKCNPYANNYIECCDDKAYIKMDIGNTIFILNQSNTV
  FDIYKPOREKIKKVRKE"
  6250..7032
  /gene="FN0900"
  6250..7032
  /gene="FN0900"
  /EC_number="3.-.-.-"
  /codon_start=1
  /transl_table=11
  /product="Metal dependent hydrolase"
  /protein_id="AAL95096.1"
  /db_xref="GI:19714463"
  /transl_table=11
  /translation="MKISILGSGSAGNSTPVEIEDYKLVDTGFSCKTEERLEKIKG
  KLSIDISALITHEHSDHNGAVIARKYDIDYITIPBEGYAGVAGVGDIGDSLNFID
  GDFITNDKVKVSPDVMGDAFRTIGFKLETONKRIASITDGIYTNVREYFQVDA
  MYTESNDPNTLNCAYPMTKERYKSNGLSNNECKPTEKMTYDLTKRYPLAHVS
  KDSNNTSLIKETLEDEFIWMIRKPCVETITDQNVTKLDFIDE"
  7042..7629
  /gene="FN0901"
  7042..7629
  /gene="FN0901"
  /EC_number="3.-.-.-"
  /codon_start=1
  /transl_table=11
  /product="DNA polymerase, bacteriophage -type"
  /protein_id="AAL95097.1"
  /db_xref="GI:19714464"
  /transl_table=11
  /translation="MDEISELMEDLKFEKAGSIGNALPKDRQEVYIIGGDMADILFY
  GNDPLVLAEDYKVSKESSGAFILRLDVLVLEPEYVITLTSREIKIKNFNEERK
  KLIDLPMOILLISPKIVVFLGKEVAQILEKKEIDPFDERGQFKKMRDIDETVLTGYV
  ETVIKARNDSCKKAAIALNPLNDKNIKERLNND"

```

gene
7619. 8164
/gene="FN0902"
CDS
7619. 8164
/gene="FN0902"
/EC_number="6.3.3.2"
/codon_start=1
/transl_table=11
/product="5-formyltetrahydrofolate cyclo-ligase"
/protein_id="AA195098.1"
/db_xref="GI:19714465"
/translation="MNKKKARLLIKERRMLSKRYIDVASDXIEELKLNDEPKNYK
TWMSYMDPKNEKTPRINTPIKMSKILLVLEKVDPKEMVITIEKNQYIYSPENKED
DSEYKSIDVITLPGVAFDRKKNVGRGRIYDFPVQPNARKIAIAFEKQIIDEG
LETDYKDKVCDLITEDIILK"
8186. 9157
/gene="FN0903"
CDS
8186. 9157
/gene="FN0903"
/db_xref="GI:19714465"
/translation="MNKKKARLLIKERRMLSKRYIDVASDXIEELKLNDEPKNYK
TWMSYMDPKNEKTPRINTPIKMSKILLVLEKVDPKEMVITIEKNQYIYSPENKED
DSEYKSIDVITLPGVAFDRKKNVGRGRIYDFPVQPNARKIAIAFEKQIIDEG
LETDYKDKVCDLITEDIILK"

Query Match 3.7%; Score 175.8; DB 1; Length 12277;
Best Local Similarity 50.0%; Pred. No. 1.5e-13;
Matches 527; Conservative 0; Mismatches 517; Indels 11; Gaps 3;

QY 627 AAAAAATGAAATCGATTATTAGTGCAGATCATTTGGCAGCAGCAATAGCAATTCAC 686
DB 12153 AATGTGTAATTAATCAGTTATTGCTTCAGAGATGGGAAATAGCAATTAATTTAT 12094
QY 687 TGTCAAGCAGATGATATGATGATTAATGAGGAGCTGATCATAGAAATATTACACATA 746
DB 12093 TACATTAATTAATGACATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 12034
QY 747 TAAACCTTACCGGAAATTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 806
DB 12033 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 11974
QY 807 ATGCAACGCAATATAGAGAGATTTATCTGACACACATACATGATTTATCTTAATA 866
DB 11973 AAGTAACTGATTAATTTAAAGAGAGCT--TGATTAATTAATTAATTTGGTGGTGGAG 11917
QY 867 TTCCCTACACATTTAGCAGCAATATGACCAATTAATTAATTAATTAATTAATTAAT 926
DB 11916 TTCCCTCAAAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 11857
QY 927 AAAAACTGCAATTAATTTAGTAAAGGATGCAATTAATTAATTAATTAATTAATTAAT 986
DB 11856 AATTAATTTGTTAGTGAAGAGCTTAAGAGATTAATTAATTAATTAATTAATTAAT 11797
QY 987 GTGAATATGACAGAGAAATTTTACATTAATTAATTAATTAATTAATTAATTAATTA 1046
DB 11796 TTAATGAGAGAGAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 11737
QY 1047 TTGCTTAAG 1106
DB 11736 ATGCAAG 11677
QY 1107 AACTGTGGAATCATGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1166
DB 11676 AACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 11617
QY 1167 AAGATATTTAGTGAAG 1226
DB 11616 CTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 11557
QY 1227 GAATTAATCGTGAAG 1286
DB 11556 GAATTAATCGTGAAG 11497
QY 1287 TGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1336
DB 11496 TAAAG 11443
QY 1347 TTGCTCAATCATGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1406

DB 11442 ATGACCTAATCGTGGTTAGAGAGATTAATCGTTACATGTCAGATATGACAGAGAGATA 11383
QY 1407 TGGCTTTTGACAG 1466
DB 11382 GAAAGAGAGAGATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 11323
QY 1467 TAAAGCTTTGAG 1526
DB 11322 ATATGAGAGAGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 11263
QY 1527 ATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1586
DB 11262 ATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 11203
QY 1587 ATTAAGCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1646
DB 11202 CAGAGAG--TGTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 11145
QY 1647 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1681
DB 11144 GTATTAAG 11110

RESULT 24
AP004827/c 333750 bp DNA linear BCT 24-APR-2003
LOCUS Staphylococcus aureus subsp. aureus MW2 DNA, complete genome,
DEFINITION Staphylococcus aureus subsp. aureus MW2 DNA, complete genome,
strain:MW2, section 6/10.
ACCESSION AP004827 BA000003
VERSION AP004827.1 GI:21204509
KEYWORDS

SOURCE
ORGANISM Staphylococcus aureus subsp. aureus MW2
Staphylococcus aureus subsp. aureus MW2
Bacteria; Firmicutes; Bacillales; Staphylococcus.

REFERENCE
1 Baba, T., Takeuchi, F., Kuroda, M., Yuzawa, H., Aoki, K., Oguchi, A.,
Nagai, Y., Iwano, N., Asano, K., Naito, T., Kuroda, H., Cui, L.,
Yamamoto, K. and Hiramatsu, K.
Genome and virulence determinants of high virulence
community-acquired MRSA
Lancet 359 (9320), 1819-1827 (2002)

TITLE
JOURNAL
MEDLINE
PUBMED
2 (bases 1 to 333750)
2 Director-General, Biotechnology Center, Aoki, K., Oguchi, A.,
Nagai, Y., Asano, K., Iwano, N., Baba, T., Kuroda, M., Hiramatsu, K. and
Kikuchi, H.

TITLE
JOURNAL
MEDLINE
PUBMED
2 (bases 1 to 333750)
2 Director-General, Biotechnology Center, Aoki, K., Oguchi, A.,
Nagai, Y., Asano, K., Iwano, N., Baba, T., Kuroda, M., Hiramatsu, K. and
Kikuchi, H.

FEATURES
source
1. 333750
/organism="Staphylococcus aureus subsp. aureus MW2"
/mol_type="genomic DNA"
/strain="MW2"
/db_xref="taxon:196620"
complement(278..619)
/gene="MW1341"
complement(278..619)
/gene="MW1341"
/note="ORFID:MW1341"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB95206.1"
/db_xref="GI:21204510"
/translation="MTNITKEVFDNLSEIDVYPAKNTLGSSEAKPYLDVSHKIIDY
FRQNDITGNDIDKLNQIPVPMNPKERYDVIKRIKHFMRGRQKTKITELIGNNA
SYOTRLKNQV"
complement(624..1283)

gene
11442 ATGACCTAATCGTGGTTAGAGAGATTAATCGTTACATGTCAGATATGACAGAGAGATA 11383
QY 1407 TGGCTTTTGACAG 1466
DB 11382 GAAAGAGAGAGATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 11323
QY 1467 TAAAGCTTTGAG 1526
DB 11322 ATATGAGAGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 11263
QY 1527 ATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1586
DB 11262 ATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 11203
QY 1587 ATTAAGCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1646
DB 11202 CAGAGAG--TGTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 11145
QY 1647 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1681
DB 11144 GTATTAAG 11110

```

CDS
    /gene="nch"
    /complement(624..1283)
    /gene="nch"
    /note="ORFID:MW1342"
    /codon_start=1
    /transl_table=11
    /product="endonuclease-like protein"
    /protein_id="BAB95207.1"
    /db_xref="GI:21204511"
    /translation="MWSKKKALMIDVIANPDAECGLKHNPELTIATVLAACCT
    GELPOTHELSLAGKRTANVWVSAPDESLAVDTHVRVSRGLKRNKQNRQ
    VEDRLCSVIFRDMNRSHQLFGRHHLAKPKCDICPLEDCRQKRYKASLKE
    A"
gene
    /complement(1273..1959)
    /gene="MW1343"
    /complement(1273..1959)
    /gene="MW1343"
    /note="ORFID:MW1343"
    /hypothetical protein, similar to chromosome replication
    initiation protein dnaD"
CDS
    /codon_start=1
    /transl_table=11
    /protein_id="BAB95208.1"
    /db_xref="GI:21204512"
    /translation="MDKIQKARPVVIRRELDHYSIDGLBODIVILKTIYSETS
    NKQPSIELQKSTQPDITWVIONLQRELELQVQDEGRFTEYMLDPPFKL
    SHLQSQMTEKQNSKEFKQLFRVLEDTPARPSPEIEITLQWIDVGDADAIO
    AALDEANSLNKLSEFKMDRILMNMKNVKTIDSRIKREKPKMTHVTKVPKD
    MNGENLDK"
gene
    /complement(2287..3579)
    /gene="aasn"
    /complement(2287..3579)
    /gene="aasn"
    /note="ORFID:MW1344"
    /codon_start=1
    /transl_table=11
    /product="asparaginyl-tRNA synthetase"
    /protein_id="BAB95209.1"
    /db_xref="GI:21204513"
    /translation="KTTIKQKDHNDVTITGAWLNKRSKGIAELQLRDGTGNG
    GVVAEVEEVEFLAKEITQSSLYVTGITTEDRSIDGYEQVSKIVASEADVP
    ITPKNGTEFTLHHRHMLRSKQKAVAKINEVIRATVEFNKDGFTKPLPILTS
    AEGTSELFTKYPDQDAFLSQSGLYLEAAMAKVSEFGTFEAKSKTRHLLIE
    PMIGEMAFTHAESELEIOEGYTHVAVKLECKLEKTLERDTKLEAVAPPR
    ISYDAIRIKSEGGPDIEMGEBGAPHEHTLAHYDLPVITVPTKIKFVQAPRP
    ENERTVLCADLIAPRGYBITIGSERVDLLEBRVKEHGLDEBASITYUDKRYGS
    VPHGFGGLSEVTAMISGVEHRETAFPRLLNLYP"
CDS
    /complement(3901..6594)
    /gene="ding"
    /complement(3901..6594)
    /gene="ding"
    /note="ORFID:MW1345"
    /codon_start=1
    /transl_table=11
    /product="probable ATP-dependent DNA helicase ding"
    /protein_id="BAB95210.1"
    /db_xref="GI:21204514"
    /translation="MGMAVAVVDLETTGNOLEPDDIIQIGTFVRRNOIITFYHMI
    RNLIEIPFIALSTIEENMLQAPYENQVNOEYDKIKDQIFVANHVDIDIANFMI
    FRCNIOYRPKVITDLEIFKIAPTDKSYOLSELBAHGHTLANHARABEDATTX
    LMITLAFKPEKLPDITLKQLYYSKQKTYDITFEFVROYDAPKPKSEKBEQIT
    TRQVDFKPTTNINGSLSLVSQVADQLGYTPQQYVLEAITLIDOLMSEKMTBA
    SLGSGKSLAYLALAMVNIETGKVMISTNKLQSQLEKDI PANMEALNFKINALL
    IKSQSDYSLGILQIDPSTNENVTIKMQLIWTETPSGDIQELNKGQKQMF
    DQLEIYVPAHDVHYNFIKRNQNIQIGTINAHLSHDSVENSIFYLQDDCTVDA
    HLPDYALNOVTNELSVADIKYQGLIGKNENELKALIDLEKORTLEKDLAPDI
    FLKASMEIHINEOLSTIETIINDVYDDIHRPHVFTRETDILKDAIITD
    KLNKLEIFNGISHTYVSLKQKLIYLDKRNATKQSKAGHTSFISKLSQKSTIR
    IYVDYAKVDVITKQVDEKFKSLFISGTLKNSFEAKQLEKNDVHNTFENSTL
    QSAKTSVFIQSVASVQYKNIDEVASIVYIILEYTTISSKCLVLFSTYKQHWQ
    DMNLPEFEDVYVLTQOQONQYKIVQGFNNDAIILGISTFEGGDFPQANGIKCM
    IAKLPFMKHNNAKWLMDSEFTSEKVEYVLPDAVTRFQGIIRLIRENDRGLIVSPD

```

```

gene
    /complement(6618..7589)
    /gene="MW1346"
    /complement(6618..7589)
    /gene="MW1346"
    /note="ORFID:MW1346"
    /codon_start=1
    /transl_table=11
    /product="truncated hypothetical protein"
    /protein_id="BAB95211.1"
    /db_xref="GI:21204515"
    /translation="MSKYSQVYLQILYKKNYISQGSIAESLNSRTIAVKVYIQLK
    LBECKIDSVSNHKGHLIQQLPDITWQGIIDQYKSSALFPSSVYDSISTOLAASKL
    VGNQSSFFLISDQETKGRGFRNHRSSSGQGLMMSVTVRPVAFSMISKPMIIFALG
    IRDAIOGFQDEVKYKPNDIYIDNGKVGCFLEWANNNGIEALICGIGINTLOLE
    NPESIRHRATSVOLHDKKIDRYOFLERLOIEKRYNGFPLTPSPSBIREYIASN
    INMRTLFTENDKQFGQALDIDYQCYILVROBAGSHRLISADIDF"
gene
    /complement(7576..8778)
    /gene="MW1347"
    /complement(7576..8778)
    /gene="MW1347"
    /note="ORFID:MW1347"
    /hypothetical protein, similar to poly(A) polymerase"
CDS
    /codon_start=1
    /transl_table=11
    /protein_id="BAB95212.1"
    /db_xref="GI:21204516"
    /translation="MDKSLFEQARPLEQIDNGFEAYVGGSVADYVGNRHIDDI
    TTSATPDEIESIFSHITI PVGKEHGTINVENENYEVTFRAEDYVDRRSGVTFV
    RDLVEDLORDPNNALAMDPAVKYDYDQGDQNNRIIRTVGAABEFQDALMI
    RCLRFQSSQPDIAETEPFAMRIOMADIKRELSIBIVLEITKMGINVERKFNLSK
    LKARNMPYFEHLDNQINVTALIDELILAIYSVKPDINYSKLEKLSNRQVQINO
    YIOINNALPSITIKREQLKRFVYDIDHLIKRWVADYKANDICGHEPLIYNLOTID
    ETLHRLPMNRKDMVNGGVLAHLNANSGLKQVLQIETAIYTVGKSNBETSLK
    WVDNHYKI"
gene
    /complement(8783..9925)
    /gene="MW1348"
    /complement(8783..9925)
    /gene="MW1348"
    /note="ORFID:MW1348"
    /hypothetical protein, similar to lipopolysaccharide
    biosynthesis-related pr homolog"
CDS
    /codon_start=1
    /transl_table=11
    /protein_id="BAB95213.1"
    /db_xref="GI:21204517"
    /translation="KKTIGTCYPSMGSGSITBELGKLAENGHEVHTSNIPIRIR
    KLPNNIFHQVAVNOYAVQYPPYDITLSTKIAEYKELGDLMLMHTAVPAIGIL
    AEMSGKDIKIMTTLHGTDITVLYGDSHQGAIKVEKESDITVTSKSLAQBTHEII
    ENKKEIIPYVRENREPEFTKNTALSKSFGIAPDEKYLIVHSNPROVORIDITIETP
    AKVREKIPSEKLLIDGPELVPRLQTELNEBEDVLFGKDDCVSEFPQLDVLILL
    SEKESFGTLLEAMKTGVVIGSNAGIKVEIKHETGTVVDVGGDSASDAIRLLE
    DVLVYKLOKXMLADIABERFSELSITDQYEVYYQKMLNHNKSKGE"
gene
    /complement(10167..10484)
    /gene="MW1349"
    /complement(10167..10484)
    /gene="MW1349"
    /note="ORFID:MW1349"
    /codon_start=1
    /transl_table=11
    /product="conserved hypothetical protein"
    /protein_id="BAB95214.1"
    /db_xref="GI:21204518"
    /translation="KMSVMEQREVEDEYIGQFKTGFSPLNLAALTEVEGLAREIN
    HTYGEKKKDSSEANTIRAEIGDNTFVLLCLANSWGIDMTSEFNETMEKFTYDRNRF
    ERK"
gene
    /complement(10820..11518)
    /gene="MW1350"
    /complement(10820..11518)
    /gene="MW1350"
    /note="ORFID:MW1350"
    /codon_start=1
    /transl_table=11

```

/product="conserved hypothetical protein"
 /protein_id="BAB95215.1"
 /db_xref="GI:21204519"
 /translation="MSIFMTIYFVILMLPLMHOKHVKNSYKSVQTSKTR
 EVALLHNGIYDVEVKGSEFLDHPKRVVSLSPAYDDBSVAGTIAHNG
 HAIDHGWTFRFRPAALPVANLSSLSYIMIGITLITGSAFSGTALMTAGLM

Query Match 3.7%; Score 174.4; DB 1; Length 333750;
 Best Local Similarity 50.9%; Pred. No. 7.1e-14;
 Matches 533; Conservative 0; Mismatches 491; Indels 24; Gaps 4;

628 AAAAAATGAAATCAGATTTTGGGAGGAGATTTGGCAGAGAAATGCAATTCGCT 687
 23025 AATGACTAAATATCCGTTTGGTATGGAGATTGGACAGCCCTTGCAGATGTTCT 22966
 688 GTGAGCAGATGATATATGATTTGAGTGGGAGCGTATGATGAAATATTAACATAT 747
 22965 TCGAGAAATGACATGATGTTTGTATGGGGTAAATATCAAGATGCTGTGATGAT 22906
 748 AATCACTTACCGAAAAATTTAAATATTTTACCCCATATCATCTACAGACAACTATA 807
 22905 AATATCATGTCATACAAATAAAAATTTTAAATATGCGGAATTAATGATTAATCAT 22846
 808 TCGAACCCAGCATATGAGAGATTTATCTGACACAAATCATGATATTTATCTTAAT 867
 22845 CGCTACTGATATGACCAAGCAATTCATTTGCAGATTTTATTAATG---CTTT 22789
 868 TCTTACACAACTATACGACCATATGATACAAATTAACACAAACGATATGTGTA 927
 22788 ACCTACTTAAAGCAAGCGAAGTGTCTTCAATTAATATATAGC-----TGACCTC 22735
 928 AATATCTCCATATTTTATTTAGTAAAGTATCGAAATTCATCATCTCAATTTCCAG 987
 22734 TAAATAGACTTTTATATCATGTTGCTAAAGTATTAATATGAGCGTTAACTGTCTC 22675
 988 TGAATATGACAGAAATTTT-----ACATATATATCCATTTTATATCTCTCG 1038
 22674 AGAAATGATGAAAGTTCATTTCACTGAAATATATGAGATTTGCGGTGTCTGAG 22615
 1039 TCCAGTTTGTGTAGAAATTTGAGAACATCTTCTGTATGATAGTCTGTGCTGTA 1098
 22614 GCCAAGTCATGCGAAGAGTGTATGTCAGAACCACTACAGTGTCTGTATCAAAA 22555
 1099 TAAATAGAACTTGTGATCATCTGATGAAACATATGATGATGTTCTTAAATAT 1158
 22554 AGATTAAGTGTATGATTAATTAAGCAAGATTTATTAATGATATTTTGGCTGTA 22495
 1159 ATACCATCAAGTATTTATGCTGACAGATTTGAGCTCATTAAGAACATTAATGCA 1218
 22494 CAGCAATGATGATGATGCTGATGCTGTAAGTGTGTCATTAATATCATTTGCA 22435
 1219 TGCATGTTGATATTCGCTGAGAAAAATTTAGTAAATATGCTGTGCTATCTATAC 1278
 22434 AGCAAGTGTATGCTGATGCTGAGTGTGCTGATGATTAATGCAAAAGTGCATTA 22375
 1279 TAAAGCATGAATGAATTTAAACCTATATATGACAAAAATCATCTCAATGATCTT 1338
 22374 TCGTGTGTTAGCGGAATTAAGTGA-----TTAGTGAAGAGTATGAGCGCAT 22321
 1339 TACATTAATGTTGCTCATGATGCTGAGATCTAATTAATCATGTCAGAAACATTC 1398
 22320 GACATTTTCTAGTTTATGCTGATGCTGATGCTGATGCTTCTGCAATCAACATTC 22261
 1399 ACGCAATATGCTTTTGAAGTAAATGAGAAAGTGAATTAATTAATCAATTAATGA 1458
 22260 TCGAATTTTCAATAGATTAATTAATCTGACAAAGTGAATCAATGATCAAGATTA 22201
 1459 TCAGAACTTAAGCTTTTGAAGAACAGATTAATTAATCAATGATTAATGA 1518
 22200 TGAATTAATATGCTGTTGAGGATTTATTAACAACTAAATCAATTAATTAATGA 22141
 1519 AATCTTATGATGATCAATTTGATATCTATTTATTAATTAATTAATGAATAT 1578

Db 22140 AGAAAAATGTCATATGTCATTAACAATGATATATATAGATATTTGAAATAT 22081
 1579 ATCAATAGATTAAGCCATATCAACATATTAATTTATGTCATATCTTGAATTTCA 1638
 22080 CTGATTAAGATTAAGTGTAAAGATTTTAATGAGCGCATTAATAATCTGAATTAAT 22021
 1639 ACAATATGATTAATTAAGCGTCAAAAT 1666
 22020 GATATTTACATATATACGCTAATAT 21993

RESULT 25
 AC005504 104992 bp DNA linear HTG 01-APR-1999
 LOCUS Plasmodium falciparum chromosome 12, *** SEQUENCING IN PROGRESS
 DEFINITION *** 3 unordered pieces.
 AC005504.3 GI:4558584
 AC005504 HTG, HTGS PHASE1.
 SOURCE Plasmodium falciparum (malaria parasite P. falciparum)
 ORGANISM Plasmodium falciparum
 Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
 1 (bases 1 to 104992)
 Hyman, R.W., Fung, E.L., Qin, F., Tanaki, T., Kurdi, O.B., Conway, A.B.
 and Davis, R.W.
 Plasmodium falciparum 3D7 chromosome 12
 Unpublished
 2 (bases 1 to 104992)
 Hyman, R.W., Qin, F., Fung, E.L., Conway, A.B. and Davis, R.W.
 Direct Submission
 JOURNAL Submitted (21-AUG-1998) Stanford DNA Sequencing and Technology
 Center, Stanford University, 855 California Avenue, Palo Alto, CA
 94304, USA

COMMENT

NOTE: This is a 'working draft' sequence. It currently
 consists of 3 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.

FEATURES

1 58642: contig of 58642 bp in length
 * 58643 58842: gap of unknown length
 * 58843 91011: contig of 32169 bp in length
 * 91012 91212: gap of unknown length
 * 91212 104992: contig of 13781 bp in length.
 Location/Qualifiers
 1. 104992
 /organism="Plasmodium falciparum"
 /mol_type="genomic DNA"
 /db_xref="taxon:5933"
 /chromosome="12"

ORIGIN

Query Match 3.7%; Score 173.8; DB 2; Length 104992;
 Best Local Similarity 46.7%; Pred. No. 1.3e-13;
 Matches 1029; Conservative 0; Mismatches 1122; Indels 52; Gaps 13;

345 AATATATAGAGTCTTATTTGACATATCAATTAATTAATTAATTAATTAATTAAT 404
 72744 AATACATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 72803
 405 ACTATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 464
 72804 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 72863
 465 AACTTAATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 524
 72864 AAAAAATATATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 72923
 525 TCATCTGACATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 584

REFERENCE 1 (bases 1 to 169546)
 AUTHORS Hyman, R.W., Fung, E.L., Qin, F., Rowley, D., Mao, J., Tamaki, T.,
 Kurd, O.B., Conway, A.B. and Davis, R.W.
 TITLE Plasmodium falciparum 3D7 chromosome 12
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 169546)
 AUTHORS Hyman, R.W., Qin, F., Fung, E.L., Conway, A.B. and Davis, R.W.
 TITLE Direct Submission
 JOURNAL Submitted (19-FEB-1998) Stanford DNA Sequencing and Technology
 Center, Stanford University, 855 California Avenue, Palo Alto, CA
 94304, USA
 COMMENT On Aug 12, 2000 this sequence version replaced gi:8810447.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 2 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

FEATURES
 Source
 1. 23466: contig of 23466 bp in length
 23467 23666: gap of unknown length
 23667 169546: contig of 145880 bp in length.
 Location/Qualifiers
 1. 169546
 /organism="Plasmodium falciparum"
 /mol_type="genomic DNA"
 /db_xref="taxon:5833"
 /chromosome="12"
 /clone="PFYAC293"
 /clone="3D7"

ORIGIN
 Query Match 3.7%; Score 173.8; DB 2; Length 169546;
 Best Local Similarity 46.7%; Pred. No. 1.1e-13;
 Matches 1029; Conservative 0; Mismatches 1122; Indels 52; Gaps 13;

QY 345 AATATATAGAGTCTTCTTTCGACATACATTTAAATCAATCAAAACAAATATAGT 404
 DB 84328 AATACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 84387
 QY 405 ACTATCTAATTAATAAAAAATCTCAATAGCTTTAGCAAGATCAAAATATATAG 464
 DB 84388 TTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 84447
 QY 465 ACACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 524
 DB 84448 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 84507
 QY 525 TCACATCTGACATGATTAACATAGTAACATTTCTTAACCTGATGATTAATTAAT 584
 DB 84508 TTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 84567
 QY 585 ATTAATTAATTAATTTGACGCTTTCTTGATTAATTAATTAATTAATTAATTA 636
 DB 84568 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 84627
 QY 637 AATGAGTATTTAGGCGCATCTTTGGCAGACATAGCAATGACCTGACGCA 696
 DB 84628 AATGAGTATTTAGGCGCATCTTTGGCAGACATAGCAATGACCTGACGCA 84687
 QY 697 TGGTATATCACTTAATTTAGGCGCATCTTTGGCAGACATAGCAATGACCTGAC 756
 DB 84688 AATGAGTATTTAGGCGCATCTTTGGCAGACATAGCAATGACCTGACGCA 84747
 QY 757 CCGAAAAATTTAAATTAATTTCCACATATCATCTACAGCAATATGCAACG 816
 DB 84748 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 84807
 QY 817 CAATATAGAGAGATTTATCTGACAGATCATGATTTATCTTAATTTCCATACA 876
 DB 84808 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 84866

QY 877 ACAATTACGACCATATGTACACAAATACACACAAACAGCATATGTAAATACTCC 936
 DB 84867 TAAATTAATTAAT-ATATATTAATTAATTAATTAATTAATTAATTAATTAATTA 84925
 QY 937 AATATTAATTTGATGTAAGGATTCGAATTTACATCTCATTAATTTCCAGTGAATAGC 996
 DB 84926 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 84985
 QY 997 AGAAGAAAT-----TTTACATTAATTCGAATTTTATCTCTGTCGCAAGTTTGC 1050
 DB 84986 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 85045
 QY 1051 TAAAGAAATTCGAACATCTCTCTGATATAGTCTGCTGCTGCTGCTGCTGCTGCTG 1109
 DB 85046 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 85105
 QY 1110 TTGGTGAATCTTGATAGAAACATTAAGTA-ATGATGTTCTAAAAAATATATACATCA 1168
 DB 85106 ATTAATTAATTAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 85165
 QY 1169 GATATTTAGGTGACGATTTGAGCTGCTTTAAGACCTTAATTCGA----TTGCATG 1224
 DB 85166 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 85225
 QY 1225 TCGAATATCGCTGGAATAAATTTAGGTATTAATGCTGCTGCTGCTGCTGCTGCTGCT 1284
 DB 85226 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 85285
 QY 1285 CATGAATGAATTAATAACACTATATATAGCAAAAAATCATTCATATGATCTCATACAT 1344
 DB 85286 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 85345
 QY 1345 AATGCTGCATCATGCTTGAGATCTTAATTAATTAATTAATTAATTAATTAATTAATTA 1404
 DB 85346 ATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 85405
 QY 1405 TATGCTTTGACCTGGAATTAAGAAAGGTAGAAATTAATTAATTAATTAATTAATTA 1464
 DB 85406 TAAATGCTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 85465
 QY 1465 CCTTAAGCTTGGAAGGAAACCGTACTGTAACCACTGATATCTGTCGCAAAAAAC 1524
 DB 85466 ATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 85525
 QY 1525 TAAATGA--GAATACCAATTTGCATATCTATTTTACATTTTATTAATGATATATATCA 1582
 DB 85526 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 85585
 QY 1583 CTAGATTAAGCCATATCAAAACATATATCTAGTCTATCATGCTTTGATTAATTCACAC 1642
 DB 85586 TAAATTAATTAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 85645
 QY 1643 AATGATTAATTAACGGTCAAAAAATATAGAAATCAACGTTCCACCAATATTAATAAGCAG 1702
 DB 85646 TAAATTAATTAATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 85698
 QY 1703 TCAATTAACCATGTTAATCACTAGATATGATTAATTAATTAATTAATTAATTAATTA 1762
 DB 85699 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 85758
 QY 1763 ATAACTTAATTAATTTCTAGCTAATTTATTAACCAATTTCTGCTGCTGCTGCTGCTG 1822
 DB 85759 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 85818
 QY 1823 CAATGCTGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1882
 DB 85819 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 85876
 QY 1883 GTTTCACACTTTTTCATACATCATGATATTAATTAATTAATTAATTAATTAATTAATTA 1942
 DB 85877 GATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 85936
 QY 1943 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2002

[illegible]

ALTHORS	Hyman,R.W., Fung,E., Conway,A., Kurdi,O., Mao,J., Miranda,M., Nakao,B., Rowley,D., Tamaki,T., Wang,F. and Davis,R.W.
TITLE	Direct Submission
JOURNAL	Submitted (29-JUN-2003) Stanford Genome Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
FEATURES	Location/Qualifiers
source	1..250421 /organism="Plasmodium falciparum 3D7" /mol_type="genomic DNA" /isolate="3D7" /db_xref="taxon:36329" /chromosome="12"
repeat_region	188..292 /rpt_type=tandem
repeat_region	240..289 /rpt_type=tandem
repeat_region	371..432 /rpt_type=tandem
repeat_region	439..527 /rpt_type=tandem
repeat_region	922..931 /rpt_type=tandem
repeat_region	951..994 /rpt_type=tandem
repeat_region	1255..1287 /rpt_type=tandem
gene	<138..>3806 /locus_tag="PFL1480w" /locus_tag="PFL1480w" /locus_tag="PFL1480w" /locus_tag="PFL1480w"
CDS	1338..3806 /locus_tag="PFL1480w" /note="match to ESTe AU087247.1, AU087385.1" /codon_start=1 /product="hypothetical protein" /protein_id="AA036382.1" /db_xref="GI:23496829"
mRNA	/translation="(KNELVFDTYHGRSINDELDYYSKKLATCSGDNVTKIDPVSLISKEVCVAELKDHSAAWKVCWSPKRYGSILASCSPFKNVITYEVINKEMEYINNEHSEVVAYIEWSPHEYGHLGCASDGTIILSYNNKSGNEGWNKSMTFAHDNGACIVAKEPYNMILLNKKLINDNNNNNNNNNNNNNNNNNNNDVINSGFKLVSGDYDQVVIWMFDNNTKEPKFYOMNDKPKHSICDVAKPMLDNPANTIIASCSDKIYLMTEDVNNNNKGGKQIKVKYNIHTKLSWPNGTILAIGCDNAIYLKENAGIEWEENCLTEENNIIODNSNMNTNNEOENHNEMNLNDNNNNNNNYTGTVYDASLYMDTNASTLLNKTPNGSNSLIPTNNNLSHGCPYQNKDPFHESCGNFPAIYNNSTQQCOOXYSNVSTGSCNNNQAGATVSNNVINSFPANNNTIQMKGPPEQMEPPEQNGPSSQNGPSPQXNGPSEQIIVPTPPPAYIWTQNSKNKNTSNTLPNNNNITPPTAPLVNNSHGASPAYASNFEQOCHGVSAIPGVPSSSPPAQOVGGSVTNNNSVSNINNKSITGCGPPLPSTISTPIIPAAGSNPFKASTSSLQOOVGHGFPPVNNRSFALNSCMPFSFHPNNNNNNNNNNNNNNNNNSGMPRPSPSOYNNNTLNPLTSVPPAPAFSDTCQGNKNSITSNFNKNNNSFNSFNHMOANLPPRPSVMNDSKPVGSIYSGGYVPHAHMNQDSNVKPNQOQYGSNTNMPRCQTHNNN")
repeat_region	1878..1943 /rpt_type=tandem
repeat_region	1884..1942 /rpt_type=tandem
repeat_region	2765..2875 /rpt_type=tandem
repeat_region	3366..3417 /rpt_type=tandem
repeat_region	3949..3978 /rpt_type=tandem
repeat_region	4114..4144 /rpt_type=tandem
repeat_region	4164..4197 /rpt_type=tandem
repeat_region	4427..4464 /rpt_type=tandem
repeat_region	4531..4558 /rpt_type=tandem
repeat_region	4560..4589 /rpt_type=tandem
repeat_region	4711..4740 /rpt_type=tandem

```

repeat_region /rpt_type=tandem 4770..4800
/rpt_type=tandem 4875..4920
repeat_region /rpt_type=tandem 5240..5289
/rpt_type=tandem 5379..5731
gene /locus_tag="PF1485w"
join(5379..5567,5647..7311)
locus_tag="PF1485w"
join(5379..5567,5647..7311)
CDS /locus_tag="PF1485w"
/codon_start=1
/product="hypothetical protein"
/protein_id="AA036383.1"
/db_xref="GI:23496830"
/translation="MSQNSDINIDENKENVNDEIKENYEFKNNMVEYCEGAYET
FKKKKKKKYIYIYIKREHYVLACISLEKVCRIHTYLNENSHSVEDEBEK
SSTGKSTVINSLRNILIDPCINILIKELRHKIYITKCNIAEKLQGCNFDQSLA
GGRLNCKCKLQENYELGKTIENNLQSLIQITKQHLVFKNELRLVKEINADI
DDNDELSQLSEITKKTLLSEKQELQKNDELINIKTLKLYNESYNDIQDYH
KSYNINISKKKRNHIFSYQNDQNNLHYENYIKGVNLANEYFNKYYDYNNK
HEYSNKGPHSKYNGKDYKDYNNMTDYNNNNNSNAYENHNKYDYNNYTY
NNKBPQYEEYRHNKDERYNNMODPKTEPOSINYPDNNSGKRSIDSRKRYR
KIDQAYKDKTKETKLDKKEHKREKTRDQREKREKRDQREKRDQREKRDQREK
KIDQREKREKREKREKREKREKREKREKREKREKREKREKREKREKREKREK
YDRKKYKXNYDEEYEDFSRGNKYNTVTSKDDMDATSEHDRVYVM"
5514..5538
/rpt_type=tandem 5538..5564
/rpt_type=tandem 5900..7007
repeat_region /rpt_type=tandem 6910..6995
/rpt_type=tandem 6914..7066
repeat_region /rpt_type=tandem 7316..7389
/rpt_type=tandem 7333..7386
repeat_region /rpt_type=tandem 7438..7489
/rpt_type=tandem 7595..7643
repeat_region /rpt_type=tandem 7801..7840
/rpt_type=tandem 7830..7863
repeat_region /rpt_type=tandem 7976..8015
/rpt_type=tandem 8179..8230
repeat_region /rpt_type=tandem 8230..8255
/rpt_type=tandem 8392..8485
repeat_region /rpt_type=tandem 8588..8639
/rpt_type=tandem 9033..9059
/rpt_type=tandem 9086..91884
locus_tag="PF1490w"
<9086..91884
locus_tag="PF1490w"
9086..11884
locus_tag="PF1490w"
/note="TMMfam hit to PF01163, RIO1/ZK632.3/MJ0444 family,
score 2.2e-55"
/product="hypothetical protein"
/protein_id="AA036384.1"

```

```

/db_xref="GI:23496831"
/translation="MEQVENYSEIKLSNDVOTSIIOGKLTNKPFRNGITPDKRAIT
NSYLDNRUTLILKALKENYNDIYGVISSGKRAYNNAKYINDELINSLKINPILGL
BKTKNTIKINNENMYNOMENHDQSDVHNSQSDERDINHFEQSDTHFTN
DKYNGFFPKNNKNGDQSGDDTNDIYNDIYNDIYDDIYNDIYNDIYNDIYNDDEK
QNFSEKEDVLEBKNNKNNYBQEPDIYKEMNDISPYDNKQSHIINILDTKRRK
GIALATKYNTSILVFKRSQYIEGFEFRNAYTKNTPRKNVQWAEKFRRLRIL
ICGLRCPYPLVRSNVTVMISMGIDNACPKKMLNFDILKMELEYIEGICILRPLF
NCLVADPSEYNLVFCNHIYIIDVSGMEHDHPSLEFKDCLINTNFEKKIGT
IYNOCPNLYTOBNQVENNTNDINRYNNTNRYEYGVNVEYGVNVEYGVNVEY
VSGINVEYGVNVEYGVNVEYGVNVEYGVNVEYGVNVEYGVNVEYGVNVEY
VPSAGYENVEILPLVLPDIYVSSSLPDIYVFLNDKKSLNVEYIYDIYDIY
KNTPIPKNLKTIQNRVYFELKRAKYVYVYFENKCKLKYKSKHOSQYAKD
EVEEQLFSMSIYPSYLNKIDIRTEKDLKJGSIYNNFISQNNKNDKQNTKE
EDPKNKEQNKQNKQNKQNKQNKQNKQNKQNKQNKQNKQNKQNKQNKQNKQNK
EHSFEENIKOYDEYKETHOELIIBQDNENEMISNNTLNILOTNSCHLSNSV
DSDTSNIDSDSCEONDEENYVSVEHEEYKFKOIIDGITRKMCLVVEQNEK
RKKIKRYOKKKKKKLT"
repeat_region /rpt_type=tandem 9636..9727
/rpt_type=tandem 10548..10711
/rpt_type=tandem 11322..11373
Query Match 3.7%; Score 173.8; DB 3; Length 250421;
Best Local Similarity 46.7%; Pred. No. 9.4e-14;
Matches 1029; Conservative 0; Mismatches 1122; Indels 52; Gaps 13;
repeat_region /rpt_type=tandem 11322..11373
345 AATATCATAGAGTCTTTATTCGACATCATATTAACAATCAACAACAATAATAGT 404
26525 AATACAAATTTAATATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 26466
405 ACTATCTAATTAATAAAAAACCTCAATGCTTTGCAAGAACCTCAATTAATGAG 464
26465 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 26406
465 ACACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 524
26405 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 26346
525 TCACATCTGACATGATTAACAATAGTAACAATTTCTTAACCTGATATATACATA 584
26345 TTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 26286
585 ATATATTAATTAATTTGACCTGGTTCTGATTAATTAATTAATTAATTAATTA 636
26285 ATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 26226
637 AATCAGTATTTAGGCGAGATCAATTTGGCAGCAGCAATGCAATGCACTGTCAGACA 696
26225 AACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 26166
697 TGGTATATACGTAACCTTAGGGGAGCGATCATGAAATTAATTAATTAATTAACCTTA 756
26165 AATGGAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 26106
757 CCGAAAAATTTAAATTAATTTACCAATATCATTCACAGCAACATATATGCAACGAG 816
26105 ATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 26046
817 CAATATAGCAGAGTATATGACAAACATACATATTAATTAATTAATTAATTAATTA 876
26045 ATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 25987
877 ACAATTAACGACATATGACAAATACATACAAACAAACGACATATGTAATAATCTCC 936
25986 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 25928
937 AATATTAATTTGATAGTAAGGTAATGCAATTTACATACCTCAATTTCCAGGAAATAC 996
25927 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 25868
997 AGAAGAAAT-----TTACAATATATATCAATTTTATATCTCTGTCAGAGTTTGC 1050

```



```

oxidase, Fe-S oxidoreductase"
/protein_id="AA035452.1"
/db_xref="GI:28203009"
/translation="MEEDSMRVEGVVPPSEANSLIIQTIGCSHNKSCFCWYK
KFRVVELKEIFEDLEQRLYYKINRIPLADGALCLTEBSLKILKIKETHPQCD
RVGIYGTADILRKSIEELKEKEGLGIYMGSGSEDEIKKILKIKVLPDEAS
KVKISGILSTVVISLIGKEKKEHAETGKAVSLIDPDLVGLVTLILNEDTMAE
QVKGERTILKPEVNIETKLMENDVNTNCFRSHASNVVPLKGTLPKDEKJLRT
IDNLSGDYYKEDERLL"
/locus_tag="CTC00855"
/locus_tag="CTC00855"
/locus_tag="CTC00855"
/locus_tag="CTC00855"
/codon_start=1
/translation="hypothetical protein"
/product="hypothetical protein"
/protein_id="AA035453.1"
/db_xref="GI:28203010"
/translation="MSBALNIIILGGFMGNANSVYKSKSLKYLSPHIVLVFLG
LVLYLKEKSNINHYIMINIMVMSFMSINIIIIKKEGENDYPRKNAIITISV
LSVVAIIMLCGLIFPKFTN"
/locus_tag="CTC00856"
/locus_tag="CTC00856"
/locus_tag="CTC00856"
/locus_tag="CTC00856"
/codon_start=1
/translation="hypothetical protein"
/product="hypothetical protein"
/protein_id="AA035454.1"
/db_xref="GI:28203011"
/translation="MEGSMNKSVYAFILMSCIISLIRQVYLAQGVSDPEQSLH
OELDLGYREGNPSSEMLVPLNIGFDKELIDYOSDANRFSKILMSGROAN
RDQVYENKNNNNLIPAPSPREGOKIESININPANIASLDGAGGAGCTPT
ALTDSAVYKAVIKSVYVGSALGVAAIMSHLSNHTGVRGKSEPYRGNTDLYLD
MYIGTRIRITY"
/locus_tag="CTC00858"
/locus_tag="CTC00858"
/locus_tag="CTC00858"
/locus_tag="CTC00858"
/codon_start=1
/product="chemotaxis sensory transducer protein"
/protein_id="AA035455.1"
/db_xref="GI:28203012"
/translation="WPKFLEBDVVLNSTLSKFORNTINVKSPAFGVGLNIMIC
LLEFVNDPEFLKEIYAVSIIELILFELKIKSKNFIIDNTVNIIGEMVITA
YNCIPENTYIPSEPCALPYEMTFCFLINKKIIITVSAIPARISEVILLILPSLI
PIGENTIMTKVYVILLINIGYILSTYANLLEKSNSEDEKIKNDMTLTFNO
IKNSITLVISKDIATSEISANLQETISACVNIOSDSWEDIKKKEITQDL
YSSNQTSDRVVTVNGSKELYSIFNKESINELISIIYSIKEGTFKPDSTKFLK
SNEIDQIAVINVSBOCTLLNLSAIEAARGAKGFVAVANEIRTLSDNTKSLQ
NIAKINDPKSLIDVENMLNNTNITGGVAILTDIVKELCTNINLNNTVNTVNEI
TEFNKLEIDIVVNLNKKISNGTELIVDEKVTVALENNTAVSEELNLSABELND
IATNNSNLIID"
/complement(4390..5400)
/locus_tag="CTC00859"
/complement(4390..5400)
/locus_tag="CTC00859"
/EC_number="2.3.1.41"
/codon_start=1
/translation="3-oxoacyl-[acyl-carrier-protein] synthase III"
/product="3-oxoacyl-[acyl-carrier-protein] synthase III"
/protein_id="AA035456.1"
/db_xref="GI:28203013"
/translation="WTKLRNVYIKSVGVYVPHKKVHNKELIQHEQYGSQNEAKELM
DEIGETISMDENETVITMSLASKICLERAGLTPNDIVIDIASNTPELSPSTGM
LIRDLNSKAVDLVNDCTGMLQGLVATSLKLNQYRALVAGFGLRSGD
NISAYVGDNSTAMILEVQBEIRFGLGHSRPSDQYIAIOFPAQGMKIPDKR
DEFEKRLPMDNNSHPIFQSSADVTETLLTYNKPDPYVTOFPGSISYENLTNVA
MSLPHNRFTPLADKGYTGPCSPPLALHDIRQEKPFNSELCIMCSVAAGCTITALLY
KW"
/locus_tag="CTC00860"

```

```

CDS
5833..6921
/locus_tag="CTC00860"
/codon_start=1
/translation="MBEYVKEKIKNTKIKFAGKAYVULALLLSIFDSNPL
SKNIIINVLIVSVRIALVGVALITGTDLISAGRVGFACISASILLQRPDVA
LFPNAQLPILIPILVIRIIGLIVINQVAVLVKVPPIATLGMVIVYGAACIYT
NAQPIGLKDFEELGTSTLGIPLNIIAGVAILVWPILINKTKGKHIVAGSNPN
AAVSGVNDLITVRYALAGALVLAARTGANNYGLMELDAIAAATIG
VSTSGGIVQGITTVLFEVNLNGLVILGSAWQVYIKGIIIGAVAFDIRKIVA
KR"
/locus_tag="CTC00861"
/locus_tag="CTC00861"
/locus_tag="CTC00861"
/locus_tag="CTC00861"
/codon_start=1
/product="galactoside transport ATP-binding protein mglA"
/protein_id="AA035458.1"
/db_xref="GI:28203015"
/translation="WVENNIILEMNGISKRFVYKALDGVDLKYKKGTYHALMGNGA
GKSTLWKCLFGIYRSDDEIVLGGKRVQFRKNADELNGISMTIHSLVPHKSVEN
VMGRPPVKKVFGIYDHKKVYEDTKDLGKLNKNIIDPTIVSKLSVQVQGLIAK
AVSYNSKIIYVDEPTSLTENEVTHFNIIISLKNQGVAIYISHMEVETLKADAVT
IMDGKIGTWEALGTTDLISKVGRLDITNRPRENTPGVIMKVEVITLANKS
FKIISEIRBELGIGTIGVAGQRTLVRSIISGLRIENGTIYNQGVYKISPIK
KXIALTERBERTGIFPLTVGDNITLNGDLYDLKVPVNOKRMKQIKNSIECLN
IRTPSHATQIKNSGNGQKIVRSMLTEPDLVIMDEPTREGIDVAKIHYISIDSL
SKWKSIMTSSMPELIGSDRIMTICDGRSLGIIIEGEPVQSEIMKATRPFI"
/locus_tag="CTC00862"
/locus_tag="CTC00862"
/locus_tag="CTC00862"
/locus_tag="CTC00862"
/codon_start=1
/translation="galactoside transport system permease protein
mglC"
/protein_id="AA035459.1"
/db_xref="GI:28203016"
/translation="MEFVYKEKIKNTKIKFAGKAYVULALLLSIFDSNPL
SKNIIINVLIVSVRIALVGVALITGTDLISAGRVGFACISASILLQRPDVA
LFPNAQLPILIPILVIRIIGLIVINQVAVLVKVPPIATLGMVIVYGAACIYT
NAQPIGLKDFEELGTSTLGIPLNIIAGVAILVWPILINKTKGKHIVAGSNPN
AAVSGVNDLITVRYALAGALVLAARTGANNYGLMELDAIAAATIG
VSTSGGIVQGITTVLFEVNLNGLVILGSAWQVYIKGIIIGAVAFDIRKIVA
KR"
/locus_tag="CTC00863"
/locus_tag="CTC00863"
/locus_tag="CTC00863"
/locus_tag="CTC00863"
/codon_start=1
/product="transcriptional regulator of the lacI family"
/protein_id="AA035460.1"
/db_xref="GI:28203017"
/translation="MMPITNDVAREGVSIISVSRVNNYVKEETRKRIEALIEKL
NKKPNMLASLITKSTIGVLPGLTFPTIVESIEDYKQGVSYVLNTRREA
LKEKEIEMLSKGVGGIIIDPTIDNKVLEIDSMKLPITIVGSEELKNGRCA
YBKVGTATAPRYLIEGHEKATFTRGSKSYVDIKKXTYDIIIRKNTKXNBIIDV
KENSIEVANTQKVEEFLNKKKPAIPACDMLAGVYVANCNKLNANVPDIDISIG
FNTLISNIIHPKLTIVDLNMRIGNIALKLEIIEDEKNSIKYMETKLIIVRESY
"
/locus_tag="CTC00864"

```


CDS

/locus_tag="CTC00864"

10734..12287

/locus_tag="CTC00864"

/EC_number="2.7.7.10"

/codon_start=1

/transl_table=11

/product="galactose-1-phosphate uridylyltransferase"

/protein_id="AA035461.1"

/db_xref="GI:28203018"

Query Match 3.7%; Score 173.4; DB 1; Length 301278;

Best Local Similarity 51.7%; Pred. No. 9,8e-14;

Matches 477; Conservative 0; Mismatches 431; Indels 15; Gaps 3;

QY 620 ATTAGGTAAATAAGTAATGAGTATTTAGGTGAGATCATTTGGCAGCAATGCA 679

DB 296990 AATATGACATTAAGGTATTAATCTTTATGAGAGAGAAATTTGTAAGCTTTAGT 297049

QY 680 ATTGCACTGTGACACATGTATATCACTTAATGAGGACGTGATCATAGAAATTT 739

DB 297050 ATTATGTTAGCCAAAAGGTTATTAATATATATGAGATAGAAAACCCCATGTGTA 297109

QY 740 AACATATTAACACTTACCGAATAATTTAAATTTTACCAATATCATCTACCAAGAC 799

DB 297110 GCGATATTAAGCAAAAAAATAATTAATTTCTCCCAATGTGTTATTCATCA 297169

QY 800 AACATATATGACACACATATAGACAGATTAATCTGACAAACATATGTAATATC 859

DB 297170 AATGTAAAGCCTATTAAGGTATGAAAGGCTTATGAGGCAAT---TAAATATGTAAT 297226

QY 860 TTAACATATCTTACCAACAATTTAGCAACATATGACAAATATAGCAACAACAGAT 919

DB 297227 ATTCTGTATCCATCTCAAGCTATAGAGAAATATGCAAAATATG-----AAAGATTAC 297280

QY 920 ATGTGTAAATAATCTCCAAATATTTATTTGTATGTAAGGTATGAAATATCATCACTG 979

DB 297281 TTTAAAGAAAGATGCTATTTATATATAGTGGCAAGGCTATAGAAACATATGAGAAA 297340

QY 980 TTCCAGTGAATATGACAGAAATTTTACCAATATATCAATTTTATCTCTCGGT 1039

DB 297341 AGGTATATCTCAATATATTAAGAGAAATTTCTTAAATATCTGTGTATATCTTTGAGA 297400

QY 1040 CCAAGTTTGTCTAAAGAAATTTGACAGACATCTTCTGTATATAGTACTTGTGAT 1099

DB 297401 CCAAGTCTATGACAGAAAGATAGCAACAGATATACCACTAGTATGATACATAGAA 297460

QY 1100 AATAAGAACTGTGGATGATGATGATAGAAACAAATATGATGATTTCTTAAATATA 1159

DB 297461 GATGTTAAGGATCTTTTGAAGTACAGAAATTTATTTAGTACAAATTAATTTAGGCTATC 297520

QY 1160 TACCATCAGATATATAGTGTATACAGATGAGGCTGATTAAGAACATATTTGCAAT 1219

DB 297521 ACTATATGATATATATAGGTGTTGAATATAGAGGGGACATGAGAAATATATGCTTTA 297580

QY 1220 GATGTGGAATATAGCTGGAATAATTTAGTATATATGCTGTGTTGCTACTGTTATACT 1279

DB 297581 GGTGTGTGATCTCGATGGAATAGATATAGAGTAATATACAAAGGCTCTTATAGACT 297640

QY 1280 AAAGCATGATATGAATTAACACTATATATAGCAAAAAATTCATATGATCTTCAT 1339

DB 297641 AAGAGAAATTAATGATATATAA-----ATTGAGAAAAGTTAGAGGAGCAAAAGGGA 297694

QY 1340 AATATTAATTTGCTCATCTGTCTTGGAGATCTAATATTAACATGTAACAAGAACATTTCA 1399

DB 297695 ACCTTTGGGGGCTAACAGGTATGGGGATATGATAGTAACCTGACTAGATGATCAAGT 297754

QY 1400 CGCAATATGCTTTTGGACTAGAAATAGGAAAGGTAGAAATATTAATATACATTAATGAT 1459

DB 297755 AGAAATAGAGAGCTGTCTTTTATATGAGAAAAGTTTATCTATGGAAGAGCTATAGAA 297814

QY 1460 CACAACCTTAAGCTTTGTAAGAGAACACAGTACTGTATTAACCACTGATATATACGAA 1519

DB 297815 GAAGTAGGATGGGTGTAGAGATATTAAGAGCTGTATTAAGCTTTATGATTAAGGAA 297874

QY 1520 AAACCTATATGTAGTACATCAAT 1542

DB 297875 AAGTTAAAGATCTATGCCAT 297897

RESULT 29

AP003134/c

LOCUS

DEFINITION

Staphylococcus aureus subsp. aureus N315 genomic DNA, complete

VERSION

AP003134.2 GI:14349226

KEYWORDS

Staphylococcus aureus subsp. aureus N315

SOURCE

Bacteria; Firmicutes; Bacillales; Staphylococcus.

ORGANISM

REFERENCE

1 Kuroda, M., Ohta, T., Uchiyama, I., Baba, T., Yuzawa, H., Kobayashi, T., Cui, L., Oguchi, A., Aoki, K., Nagai, Y., Lian, J., Ito, T., Kanamori, M., Matsumaru, H., Maruyama, A., Murekami, H., Hosoyama, A., Matsumaru, U., Y., Takahashi, N. K., Sawano, T., Inoue, R., Kato, C., Sekimizu, K., Hirakawa, H., Kohara, S., Goto, S., Yabuzaki, J., Kanehisa, M., Yamashita, A., Oshima, K., Furuya, K., Yoshino, C., Shiba, T., Hattori, M., Ogasawara, N., Hayashi, H. and Hirakawa, K. Whole genome sequencing of methicillin-resistant *Staphylococcus aureus* Lancet 357 (9264), 1225-1240 (2001)

JOURNAL MEDLINE 21311952

REFERENCE 11418146

2 (bases 1 to 301550)

Director-General, Biotechnology Center, Aoki, K., Oguchi, A., Hosoyama, A., Nagai, Y., Kuroda, M., Hirakawa, K. and Kanamori, H. Direct Submission

TITLE

Submitted (30-JAN-2001) Director-General, Biotechnology Center, National Institute of Technology and Evaluation, Biotechnology Center; 2Chome 49-10 Nishihara, Shibuya-ku, Tokyo 151-0066, Japan (E-mail: bio@nite.go.jp, URL: http://www.bio.nite.go.jp/, Tel: 81-3-3481-1933, Fax: 81-3-3481-8424)

COMMENT

On Jun 12, 2001 this sequence version replaced gi:13701258.

FEATURES

Location/Qualifiers

1..301550

/organism="Staphylococcus aureus subsp. aureus N315"

/mol_type="genomic DNA"

/strain="N315"

/sub_species="aureus N315"

/db_xref="taxon:158879"

complement(249..947)

/gene="SA1293"

complement(249..947)

/gene="SA1293"

/note="ORFID:SA1293"

/codon_start=1

/transl_table=11

/product="conserved hypothetical protein"

/protein_id="BA042553.1"

/db_xref="GI:13701259"

/translation="MSSIFTMIIYFVILMTVLEPLMAOHKVKSNYKYSQVSTGKTR EYALHILHANGIYDVEYVKGEGFLDHYDPKKVLSLPANVDRSPVAGTAAAEVGH HAIOHOGWFLRFPALVPAVNLGSSLSYIMIGILITAGSFGSTALMIGAGLM SLAVFSYITLVEFDASSRAMKQITALNINVEKEXKARVLSAAATVYAAATVAV AEVILILIRSSD"

complement(1001..1588)

/gene="SA1294"

complement(1001..1588)

/gene="SA1294"

/note="ORFID:SA1294"

/codon_start=1

[illegible]

Query Match	Best Local Similarity	Score	Length
Matches 532; Conservative	50.8%;	172.8;	301550;
	0;	Mismatches 492;	Indels 24; Gaps 4;

628 AAAAATGAAATACGATTTTACGTGACAGATCATTTGGACAGCAATACCAATGTGACT 687

```

Db      12452 AATGCTAAATTTACCCCTTTTGGATGGAAGTTTGGGACACCCCTTGGCAATGTTCT 12393
Oy      6688 GTACACACATGATATATCAGTTAACTTATGGGAGCGTATCATGAAATATATACATAT 747
Db      12392 TGCAGAAAATGACATGATGTTTGTATGTGGGATTAATAATCAAGTCTGTATGATAT 12333
Oy      748 AAACATTAACGAAAAATTTAAATATTTTACCAATATCATCTACCAACAAATATA 807
Db      12332 AAATCATGCTATACAAATTAATAAGTATTTAAATACCGAAATTAAGTATTAACATCAT 12273
Oy      808 TGCAACCGCAATATAGACGATATATCTGACAAACATCATGATATATCTTAATAT 867
Db      12272 CCGTCTTCAGATATGACCAAAAGCATTCATTTGCAATATTTACTTAATG---CTTT 12216
Oy      868 TCCTACACAAATTAACGACCATATGTCACAAATACAAACAAAGCATATGTGTAA 927
Db      12215 ACCCTAAAGCAATGCGAAGTGTCTACTCAATTAATGATAGC-----TGACCTC 12162
Oy      928 AAATCTCAATATTTTGTAGTAAGGTATGGAATTCATCCTCAATTTCCAG 987
Db      12161 TAAAAAGCTTTTATACATGTTGCTTAAAGGATTAAGAAATGCGATTAACGCTGTC 12102
Oy      988 TGAATAGCAGAGAAATTTT-----ACAAATATATCAATTTTATCTCTCTGG 1038
Db      12101 AGAATGATGTAAGTATCTATTTACCGCAATATATGCAAGTATGGCGTGTGACAG 12042
Oy      1039 TCCAACTTTGCTTAAAGAAATGCGAAACATCTCTCTGTATGATGATCTGCTGTA 1098
Db      12041 GCGAAGTACGCGAAGAGTGTAGTCACCAACCACTACAGTGTGTCATATCAAA 11982
Oy      1099 TATAAAGCACTTGTGATCATCTATAGAAACATAAGTATGATGTTCTAAATATAT 1158
Db      11981 AGATTAAGTATTAATTAATTAACGCAAGTTTATTTATGATGATTTATTTGGGTGTA 11922
Oy      1158 ATACCATCAAGATTAATTAAGTGTACAGATGAGCTGCAATTAAGAAATTAATGCAAT 1218
Db      11921 CACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 11862
Oy      1219 TGCATGGAATTAATCGCTGGAATAAATTTAGTATTAATGCTGTTGCTATCTTTATAC 1278
Db      11861 AGCAAGTGTATCTAGCTGGAATGCGCTATGATGATGATGATGATGATGATGATGAT 11802
Oy      1279 TAAAGCATGAATGAATTAATAAATCTATATATAGCAAAATATCAATTAATGATCTTCA 1338
Db      11801 TCGTGTATGCGGAAATTAAGTGA-----TTAGTGAATAAGTATGCGGATCTAT 11748
Oy      1339 TACATTAATGTCATCATCTGCTGAGATCTATATTAATGATGATCAACAGAAATTC 1398
Db      11747 GACATTTCTAGTGTAGTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 11688
Oy      1399 AGCAATATGCTTTTGAATGAAATAGAAAGTGAATAATTAATTAATTAATTAATTAAT 1458
Db      11687 TCGAATTTTCAATTAAGATTAATACTGCAAGATGATGATGATGATGATGATGATGAT 11628
Oy      1458 TCAACAACCTAAAGCTTTGTAAGAAACAGTATCTTAATAACATGATTAATGACAA 1518
Db      11627 TAAATGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 11568
Oy      1519 AAAATCTATAGTGAATCACTCAATTTGATATCTATTAATTAATTAATTAATTAATTAAT 1578
Db      11567 AAAAAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 11508
Oy      1579 ATCACTAGTAAAGCATATCAACATATTAATGATGATGATGATGATGATGATGATGATGAT 1638
Db      11507 CTCAGTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 11448
Oy      1639 ACAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1666
Db      11447 GTATATTAACATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 11420

```

RESULT 30

```

AE015938/c
LOCUS      AE015938      303414 bp      DNA      linear      BCT 06-FEB-2003
DEFINITION Clostridium tetani B88, section 3 of 10 of the complete genome.
ACCESSION AE015938 AE015927
VERSION    AE015938.1 GI:28202752
KEYWORDS
SOURCE
ORGANISM   Clostridium tetani B88
            Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
            Clostridium.
REFERENCE
AUTHORS    1 (bases 1 to 303414)
            Bruggemann, H., Baumer, S., Fricke, W.F., Wierer, A., Liesegang, H.,
            Decker, I., Herzberg, C., Martinez-Arias, R., Merkl, R., Henne, A. and
            Gottschalk, G.
TITLE       The genome sequence of Clostridium tetani, the causative agent of
            tetanus disease
JOURNAL     Proc. Natl. Acad. Sci. U.S.A. 100 (3), 1316-1321 (2003)
MEDLINE     22457253
PUBMED     12552129
REFERENCE
AUTHORS    2 (bases 1 to 303414)
            Bruggemann, H., Baumer, S., Fricke, W.F., Wierer, A., Liesegang, H.,
            Decker, I., Herzberg, C., Martinez-Arias, R., Merkl, R., Henne, A. and
            Gottschalk, G.
TITLE       Direct Submission
JOURNAL     Submitted (31-OCT-2002) Goettingen Genomics Laboratory, Institute
            of Microbiology and Genetics, Georg-August University,
            Griesbachtstr. 8, Goettingen 37077, Germany
FEATURES
Source
            1..303414 "Clostridium tetani B88"
            /organism="Clostridium tetani B88"
            /mol_type="genomic DNA"
            /strain="Massachusetts"
            /sub_strain="B88"
            /db_xref="taxon:212717"
            88..1674
            /locus_tag="CTC00580"
            88..1674
            /locus_tag="CTC00580"
            /codon_start=1
            /transl_table=1
            /product="ABC transporter ATP-binding protein"
            /protein_id="AA035197.1"
            /db_xref="GI:28202753"
            /translation="MTSTNLSLRGGRVLEPDVNVKFTPNQCYGLIANGAKSTPL
            KILSGEIEANKGDVILEPGRKVAIVKQHPQDSFVLETVIWEKRLNINKEKEXI
            YKEDFSEEDGKVALEBGEFELNIGWESAEALTLGLGIVELHKLKXDLGQ
            KYRVLLAQPLFKGPDIILLDEPTNLDIKSIOWLEFLINERTVILISHRHFANKY
            CTHLDLDFGKIOLYVGVNDFEYESSOLAEVAKQKKKEKIDLEEFARFANA
            SKSKATSRKKMLDKITLEDIKPSRRRPYVFGKEDREGNDLLAVEGKTIIDEXY
            LNNVSRVYKQKIAFVGENSIKAKRLSOLLGENDKADSGEPKGIITTSQSYFEDNS
            NYFNGVNLVLVWMLKQFSEKSEETIRGFLGMLSGEALKAASAVISGGEVRCMLS
            KMLSGANVLLIDEPNHLDESTRVANNGLINPFTILLFSDHQFQVTIANRIIEI
            TPDGIIDRQMTYDELYNSKR"
            complement (1752..2954)
            /locus_tag="CTC00581"
            complement (1752..2954)
            /locus_tag="CTC00581"
            /codon_start=1
            /transl_table=1
            /product="predicted amidohydrolyase"
            /protein_id="AA035198.1"
            /db_xref="GI:28202754"
            /translation="MKERRARSLIIRKQNLINMADIYEERDILENGKILIEISNEI
            DISKHDYKILDAKNGFVYRPGIVEPHECOLGIEEIHRRPDGDSNEEDPILPOLAID
            AINPDDEGDAKRAKQVTVITGPRESNILGTFPAVYGTGVDWMTKDLARFES
            LGNTVYNTIGKKNMPKTRWASAMITRDLIAKAYHRLYQSLINDKXKUNPREDI
            KLSLNRVYDGLVRFTHQAVDILTAIRIGEEFLANTVDCSEAYLIPRLKYSTI
            KCVIGPVGKRDPEVRNDDIIGAIMENNDNFAISSGHPVNLIELTVQITMYKK
            GSRKALKAAMTINAKITIGLEDRIGSIEIGDADIITWNGEPIDYYSYPTVLIDGN
            VVYKKEVI"
            complement (2908..4116)
            /locus_tag="CTC00582"
            complement (2908..4116)

```

```
/locus_tag="CTC00582"
/codon_start=1
/transl_table=11
/product="predicted amidohydrolase"
/protein_id="AA035199.1"
/db_xref="GI:28202755"
/translation="MLIKNCLIDMVGIIHKEKYLILNKKRIKRIEIDEKOLENV
OVDAKGRIVTEPLIDPHCHGVGTATLEQDNGNETSPSTPELRAIDAIDPVDVY
DVAIGHVTVVTPGDSANVIGGFTAKTNGNIKDCVKEVCMAKALGMEKVPY
GKQKAPKPRMNNALIMRENLEPKAEVHSKKKEKKNESDKNISEFKYDIGNMSLR
VVDGRVAVNAQSDIDITAIRICEPELRTIHCRTGYIYDELKSKVQCICIGPT
AGGSKLEIQHNEKICARTFESGIDFALMTDNPVITPEGLMQLAFVKNLSKETPA
LKITINARLITDIDRVGSIIEIGKDADIVIMVDPIDTSEAGIVIIDGVYVERA
GRIIDVYKLL"
complement(4188..5753)
/gene="ntrP"
/locus_tag="CTC00584"
/complement(4188..5753)
/gene="ntrP"
/locus_tag="CTC00584"
/codon_start=1
/transl_table=11
/product="putative efflux pump component ntrP"
/protein_id="AA035200.1"
/db_xref="GI:28202756"
/translation="NAKSKRSKSLFDFRIQGVAVAGNKLPTPELFLVLIILICVLSV
LSMGVSVYMAAGKQKAMKETTAAVAVNLISKEQPSKPLDFTKTEPEPLDLMV
NMGGVETGFTINAKRKTLSAPSLITAVAFVGNANIASDAMITPATIGAI
LEKSGRNPIYIGITIGPAAGSGGFTANFVAGTDALISGISESAEABARITPHPI
NYPMAKAPFIVTGVVTLIAEKYIIPKIDSGSGEDCDLAHELTDLEKGLSNAGI
ATITITITLAIKPKNGLRADDTITPKPSLISGIVPLFFPFLTYGIVGTTGY
ITSEKDIKPKMNSLSGALPLVAVLPSITIKFPEKSKLTVAVKSGSELKRNAG
VPLPLSFLVALVNMFTSGSSKMLLAPIFPMFEMNLSPAVSQVARVDSAT
NIISSISLAVIGIMEQYNEKDEPVIGITVALTLFPALGILLSKSLILIMVTE
DPLGPNIFLS"
/locus_tag="CTC00585"
/locus_tag="CTC00585"
/locus_tag="CTC00585"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AA035201.1"
/db_xref="GI:28202757"
/translation="WKGKTIACKLKEEDIRORLYEBSLKKEISNVKHNPIIDPEA
NKSXVPAETTFNNDCVKVESNTVALEDDDEDTIAELPIATVAVFELDEEG
TLSEKKEIPYKDIYCHLSIDIRQDSKYEGLRPLWGINQFVIGGLINNGKIMSNFQDI
VEDISNKEK"
/locus_tag="CTC00586"
/locus_tag="CTC00586"
/locus_tag="CTC00586"
/codon_start=1
/transl_table=11
/product="hemolysin II"
/protein_id="AA035202.1"
/db_xref="GI:28202758"
/translation="WGTNMFYRKKEEINAIHTGIGVPSIYAVLVAVPAKTKYKDAW
YVTSISIGSTILILMCSSTLYHSTNNEKVKIKFPPHSSIFLIIACTYPTPTITL
RQMLGWSIIGIIVTIVIGIVKIVCFEEMERKSTFLLIAGMWIVVAKSIISSLPV
KGIIVLIAGLIIYVGCIFVAKDKIPLYNAIWHFVLVGSVCHPFSILLY"
/locus_tag="CTC00587"
/locus_tag="CTC00587"
/locus_tag="CTC00587"
/codon_start=1
/transl_table=11
/product="conserved protein"
/protein_id="AA035203.1"
/db_xref="GI:28202759"
/translation="MEVYISPPVYTFYTFYIKFISGCLIMLNLIELIADIIRKGA
YAKDPLEVEHNNKREYCKENLGHVGTNNMCPGGEIEDLKEKVMNKKGIYIO
TTHPVNTRIDRAETVIRDKHNNFTLELLENREKIIDLEFFPMAGCELCCTTCGYL
```

```
KEEKICYPQKAISAEAVGINLSLKSCLQFNYSDDSLAYVGLIK"
complement(8089..9645)
/locus_tag="CTC00588"
/complement(8089..9645)
/locus_tag="CTC00588"
/codon_start=1
/transl_table=11
/product="ABC transporter ATP-binding protein"
/protein_id="AA035204.1"
/db_xref="GI:28202760"
/translation="MSIISVNTLTHGFDRALFNDVSENLKGEHIGLVANGEGKSS
FNATITGNLRPBGVIVEMSKRVYGVLDQHTVILKSGTTIRELTKTAPKLPLEKEMV
ETDKKGVANSDELELLEDSGIIQILTNDFPIIDAKIDVAVAGKAGADIGLNDV
QDLSGGQRTKILAKILLERPDILLDEPTNLDQKLEMLKRYIQDENATILSHD
IPELNSVINLIYHMEGELNRYVGVYDPLKYAEAKQDLEAAYKQDQELQELQFV
SRKARVSTRNAMSROKLDLDMDLIELPKDKPEKPFNFKEGTSGKMLPETKIDV
YNEALSPLNLHMERGETALVANGALVTLKSLGIRKIPISGALTGDVINTGYF
EDEIKYANNCTIEYWSFSPFQYEVAAALAKGLTTRKHSKVLVSGGEQAKVR
LCLINKETNPVLDEPTNHLVDVDAKEBLKRAIKAYKGSVLLICHEPEFYQDIIVDVW
NCENWTKIV"
/locus_tag="CTC00589"
/locus_tag="CTC00589"
/locus_tag="CTC00589"
/codon_start=1
/transl_table=11
/product="conserved protein"
/protein_id="AA035205.1"
/db_xref="GI:28202761"
/translation="MDKNIEWYIDKKVCRITENLEKNINAEAVYDEDELKKNIDEMI
PKGVITAGDSMTLELTVIGIDLRNDKYIYLDKHEGITESEKNTYNSFTAHVYLS
STVALTESEGLVINDGSRVAMVYGPQVIVAGINKIVNLEAEKRVNVAFL
DARKLRNPKCKIGICVDCKSSERCNFTYIKQFKRITVITNKLGV"
complement(11064..11065)
/locus_tag="CTC00590"
/complement(11064..11065)
/locus_tag="CTC00590"
/locus_tag="CTC00590"
/BC_number="1.15.1.1"
/codon_start=1
/transl_table=11
/product="superoxide dismutase (Cu-Zn)"
/protein_id="AA035206.1"
/db_xref="GI:28202762"
/translation="MAFAVCAIKGGLAPISIRGVYFTDPGGCKVCTINGLIPYKP
ASGNMPPVGHGFHIEFGNINIDPKDFKSGMGNPTNPHGADFPVIFSN
GYSFMRFTNPKVYKDIILKSIIIHSPDYNTPQPGDSGRRLACVIGRV"
complement(11185..12345)
/locus_tag="CTC00591"
/complement(11185..12345)
/locus_tag="CTC00591"
/locus_tag="CTC00591"
/codon_start=1
/transl_table=11
/product="conserved protein"
/protein_id="AA035207.1"
/db_xref="GI:28202763"
Query Match 3.6%; Score 172.8; DB 1; Length 303414;
Best Local Similarity 49.1%; Pred. No. 1.2e-13;
Matches 517; Conservative 0; Mismatches 532; Indels 5; Gaps 2;
QY 2395 TCTAATAATCATATATCTTTTAAAGATGATCTTACATATCTATCAATAATTA 2454
DB 127005 TTTGAATTAACAGATMAAAAGTGTGTTTCTTACATTTGACTTCTAATGTGTT 126946
QY 2455 TACATGCAATGTTTATTCAGAAACATTCACATCAAGATTTTCAATTTAGCGGTAA 2514
DB 126945 TTTTGTACTTGTTCCTTACATCTTTACATATATTTGAATTTCAAAATTCATATTA 126886
QY 2515 TGTGAAAGTTTCATTCGAACAAATTCAGAGTTTCGATACATTTACCTGCAAAAGCA 2572
DB 126885 TGTTTACCTATCTTTTCATCAAGAGTTCTATATATATGAGCTTCTTAAAT 126926
QY 2573 GATACAGAGTTTAAATATCTGCTTCAAAATCTTCAAGACCTTACGTAATGTATCA 2632
```

TITLE	COMMENT
JOURNAL MEDLINE	2311952
REFERENCE	11418146
AUTHORS	2 (bases 1 to 346900)
TITLE	Ohta, T.
JOURNAL	Direct Submission
	Submitted (28-FEB-2001) Toshiko Ohta, University of Tsukuba College of Medical Technology and Nursing, Department of Medical Technology, 1-1-1 Ten-nodai, Tsukuba, Ibaraki 305-8577, Japan [E-mail: tohta@tsukuba.cc.tsukuba.ac.jp, Tel:81-298-53-3454, Fax:81-298-53-3454]
FEATURES	On May 29, 2001 this sequence version replaced g1:13875626.
source	location/Qualifiers
	1..346900
	/organism="Staphylococcus aureus subsp. aureus Mu50"
	/mol_type="genomic DNA"
	/strain="Mu50"
	/sub_species="aureus Mu50"
	/db_xref="taxon:158878"
gene	12..263
	/gene="SAV1313"
CDS	12..263
	/gene="SAV1313"
	/codon_start=1
	/transl_table=1
	/product="hypothetical protein"
	/protein_id="BAB57475.1"
	/db_xref="GI:14247084"
	/translation="MLTILTKKEDQITTYEYIPEDDISNGKGSVTFNKKDAEVIDSSLSIEIYDFMLYRNKSPSVVRDFIEKQEPENYKIAY"
gene	987..1181
	/gene="SAV1314"
CDS	987..1181
	/gene="SAV1314"
	/codon_start=1
	/transl_table=1
	/product="hypothetical protein"
	/protein_id="BAB57476.1"
	/db_xref="GI:14247085"
	/translation="NNIDISYIGLVKLTNNKILIGKVIDFDKVNDFGVNSIEIDTGRITVDISENKRITVLOK"
gene	complement(1725..2750)
	/gene="SAV1315"
CDS	complement(1725..2750)
	/gene="SAV1315"
	/codon_start=1
	/transl_table=1
	/product="conserved hypothetical protein"
	/protein_id="BAB57477.1"
	/db_xref="GI:14247086"
	/translation="WISFENDYLEGAEHKLKLVDTNLVQASGCGDQFLQAIEKIKDITDPNATIRPLVGTQTNQVIVNSMSSYSGVLSADGHAVHGGAIEYSGHKKRITIPSEKGVASADVEYMETPEKSPDFKDMPPGVAIISHPLTEYGLYSSELEELCKVKQKQOLP.FMDGARLGYGVMDSQDMTKDLIAKYGDVYIGTKALCGEALVFPLAKKQNEKQPTFRIRKHGGLAKGRLTGIOPLFDNLNYFISRAIEMANMKQGFKKKGRIRLPDSPTNOQFLISNEKIALEEQKFAVMEKCYDDQHRVVRATSMATTEENL
gene	3020..3217
	/gene="SAV1316"
CDS	3020..3217
	/gene="SAV1316"
	/codon_start=1
	/transl_table=1
	/product="hypothetical protein"
	/protein_id="BAB57478.1"

```

/db xref="GI:14247087"
/translation="MEQIKLKTFTAEELLEKNINMFLSSEATNKLVTITKEIE
ERTPNNEEFPAITLSVNK"
/gene="SAV1317"
3238..4755
/gene="SAV1317"
/codon_start=1
/translation="cardiolipin synthetase homolog"
/product_id="BAB57479.1"
/db xref="GI:14247088"
/translation="MFVFPFGKTMKPTPSNDLGLFTLLAGFTINLVAFII
FLERRKSTASFWMLFVLPVPLIGLILPFGRTVSARKLAKNNGVLDLPGGLK
QOIESFDKNGETDNQVQKHDLVSKLMDQDELLENKVDH IDGNLDYQVLD
IKNAEXYHLEHYTPALDGLKRIHLAEKLGKQLEVKILYDVGSKVMANFDF
KSLGGEVAFASKLPLNFRNNNNHRIKIVIDQGLYGVFNIGDEYLGW
RDLRIQGDVADALQRLFLDMNSQAHKPOEYDVKKFKNGPLGNSPGLAAGP
ASDHQIEYGYTKMTMSAKSVLYOSPYEIPDNSYINAKIAKSGVDVHLMTCRKP
HPLVYATFNSADSLSSGVKITYTENGFHSKGLIDDELIVSGTANMPREELN
ENAFVDEMLAKDLRAVYEHDTIKSKQITKESTANRPLSVKFEKSLAKVSPIL"
4939..5436
/gene="SAV1318"
4939..5436
/gene="SAV1318"
/note="ATP-binding protein"
/codon_start=1
/translation="ABC transporter homolog"
/product_id="BAB57480.1"
/db xref="GI:14247089"
/translation="WIOSNINKSPNRCVAKNISPDIEGKCIALIKGKAGASTLI
DILIGNVANGSEIFDKDLQSENRSTIMPKMTXPKQVLEIINLYGSEYENPL
EELIETLTPDSQNLQFNKLSGGQRLDLVSLIGQPOLILDEPTSTMDIEIKY
FMSII"
5461..5838
/gene="SAV1319"
5461..5838
/gene="SAV1319"
/note="ATP-binding protein"
/codon_start=1
/translation="ABC transporter homolog"
/product_id="BAB57481.1"
/db xref="GI:14247090"
/translation="WLTSHYIEEYERYSKIIILENGEIIINDSTHIRTNOOST
LSDEYRKLKDKDLVYQKHNNGTITKITSNVDITLYLOQLHINDDIEIQNSIV
DSIFNKKQKQSGNSIDTKLENRI"
5807..6538
/gene="SAV1320"
5807..6538
/gene="SAV1320"
/note="similar to ABC transporter integral membrane
protein"
/codon_start=1
/translation="hypothetical protein"
/product_id="BAB57482.1"
/db xref="GI:14247091"
/translation="WLSYKIEBFVIRKKTLLISLFPVIFYLFTSLIEPEV
KPKFYKEMYSVTYSLISLSLTPDIDINEKONEMQRQMTPEFTSYKSVK
TMLQRAIILVFNVGHFTKGVASAVOMLESGLFWMGASLLTFTGLBSLNDIK
TSLANIVITGLAVIGLWFPINTFPNLQHVAVLBSYHLRKLGVDIASHEHNLIS
FALLVYLGSIAYVCISHFKRAE"
6542..7633
/gene="SAV1321"
6542..7633
/gene="SAV1321"
/note="similar to two-component sensor histidine kinase"
/codon_start=1
/translation="hypothetical protein"
/product_id="BAB57483.1"

```

```

/db xref="GI:14247092"
/translation="MKFLKDSIAEISSILYIFPIAGIFENEVYGPKNLYISLVF
SLSYLILVNNRINTLMFYILLIHYFICVFSVSHMLSLFVSAFAVPEIKN
NKKATKINLEILMTCITITLYLVNNAVAVVYVSLIMDNKSKRREKQK
IAEKRNHTLLIAEOERHIGODLHDITGLSISSELYKILIDTVEKVAELL
AINKSRSESLNKRZLIDVYKPSFIEZDSIRKILKADIDPFTENRELAOVSPTK
OSMLVMTREAINNVIRKNAKSKVHGKLTNNHKLMLTIEDGKIGIDCEKSLISQ
RVQHLNGTLVADSTNGTKIIEISTGILA"
7630..8232
/gene="SAV1322"
7630..8232
/gene="SAV1322"
/note="similar to two-component response regulator"
/codon_start=1
/translation="hypothetical protein"
/product_id="BAB57484.1"
/db xref="GI:14247093"
/translation="WTSUIIADQWMLRQAVOLIKHGDPEILDNTDGLDMLKLE
EYRNNAVILDIEMPCMTGLEVLAEIRKHLNIVITTPFRPGYPERAYVVDYAV
LKRSTIEELVETINKVNGEKSATLTSFVDQPLTFKQIVLREIGNLSKREI
SEKLFITDGTVRNYSVLIIDKLPADNRFDMARKANEKMTI"
8233..8533
/gene="SAV1323"
8233..8533
/gene="SAV1323"
/complement(8345..8533)
/gene="SAV1323"
/codon_start=1
/translation="hypothetical protein"
/product_id="BAB57485.1"
/db xref="GI:14247094"
/translation="WNISNENVTLLAVIAIIIGFIQDIFPKLPLIVAVLSILGI
FVGFITVLYVSPNKRKN"
8672..9205
/gene="nuc"
8672..9205
/gene="nuc"
/note="SAV1324"
/codon_start=1
/translation="therm nuclease"
/product_id="BAB57486.1"
/db xref="GI:14247095"

```

Query Match 3.6%; Score 172.8; DB 1; Length 34690;
Best Local Similarity 50.8%; Pred. No. 1,1e-13;
Matches 532; Conservative 0; Mismatches 492; Indels 24; Gaps 4;

CY	628	AAAAATGAAATGATGATTTTACGTCAGATCATTTGGCAGCAGCATATGCAATTCGACT	687
DB	199484	AATGACTAAATTAACGCTTTTGTATGGAAAGTTTGGGACAGCCCTTGCCAAATGTTCT	199425
CY	688	GTGAGCAGATGGAATACGTTAATCACTTATGGGAGCGTGATCATTAGAAATATTACACATAT	747
DB	199424	TGCGAATAATGACATGATGTTTGAATGCGGTAATAATCAAGATGCTGTTGATGAAT	199365
CY	748	AAACACTTACCGAAAAATTTAAATATTACCCACATATCTTACGACACATATAT	807
DB	199364	AAATACATGATACAAATAAAGATTTAAATACGGAATATAGATTTAACATCAT	199305
CY	808	TGCAACACAGATATAGAGAGATATATCTGACACACATATACATATATCTTAATAT	867
DB	199304	CGCTACTTCAGATATACCAAGACATTAATTTGCAAGATATTTACTTAATG--CTTT	199248
CY	868	TCCTACACAACTTACGACCATATGATACACAAATATCAACAAACAGCATATGTGTA	927
DB	199247	ACCTACTTAAGCAATCGAAGATGCTACTCAATATATGATTAAGC-----TGACCTC	199194
CY	928	AAATACTCCAAATTTAATTTGATTAAGTATCGAATTTACATCATCTCAATTTCCCG	987
DB	199193	TAAAAAGACTTTATACATGATGCTTAAGATATGAAAAAGGACGTTTAAACGTGTCTC	199134
CY	988	TGAATATGACAGAAATTTT-----ACAATATTAATCAATTTTATATCTCTCG	1038

FEATURE SOURCE COMMENT
 Direct Submission
 Submitted (09-FEB-2000) Submitted on behalf of the Campylobacter sequencing team, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk
 Notes:
 Details of C. jejuni sequencing at the Sanger Centre are available on the World Wide Web.
 (URL, http://www.sanger.ac.uk/Projects/C_jejuni/).
 Location/Qualifiers
 1..282183
 /organism="Campylobacter jejuni subsp. jejuni NCTC 11168"
 /mol_type="genomic DNA"
 /strain="NCTC 11168"
 /db_xref="taxon:192222"
 complement(43..1443)
 /gene="Cj1009c"
 complement(43..1443)
 /gene="Cj1009c"
 /note="Cj1009c, unknown, len: 466 aa; 29.1% identity to Hp0282"
 /codon_start=1
 /transl_table=1
 /product="hypothetical protein Cj1009c"
 /protein_id="CA873265.1"
 /db_xref="GI:6968445"
 /db_xref="GOA:Q9PNT1"
 /db_xref="SPTRMBL:Q9PNT1"
 /translation="MNNILIIIDGILAKPLERLCPEKGLGYFETVVCQSEKNNLNI
 SFGVLSYNDHPDPTSTALRIMSKDQKQAFVWODEPTEKSYEALRSLNPIEIM
 DPNGLSYNDHPALADARNTLSRRPMDPDLITAOYLIGLKEVNPISGIFA
 YRILISIIQQKMRIVLIRNSKIFYFVPSVLEPNSILVGPVVLQSI.FHNRGA
 GQPPMPSGNSVAFALIDMKNNQMGREVDLTILKIQKSAKFF.HVNPRLGVME
 KLKRLSDKGVGFEDYFNTDFKQISWLNQNDGLVYTDLKNFEKQAFD.KIPI
 KVEASEDLKEALII.SADESENNANNVTTDSKQDEFVILVYVYVNSQNTDME
 YFSLSLGLYDNIQIINKDENPFLNLQYREDLLQFVS.FKELLNRDFA.NLSNLR
 HYKKNQNYGLTFPFE"
 1473..12584
 /gene="tsg"
 1473..12594
 /gene="tsg"
 /EC_number="2.4.2.29"
 /note="Cj1010, tsg, probable guanine
 tRNA-ribosyltransferase, len: 373 aa; similar to many e.g.
 TGT_AECOL1 guanine tRNA-ribosyltransferase (EC 2.4.2.29)
 (375 aa), fasta scores; opt: 1016 z-score: 1190.9 E(): 0,
 41.3% identity in 373 aa overlap. 56.7% identity to
 Hp0281"
 /codon_start=1
 /transl_table=1
 /product="guanine tRNA-ribosyltransferase"
 /protein_id="CA873266.1"
 /db_xref="GI:6968446"
 /db_xref="GOA:Q9PNT0"
 /db_xref="SWISS-PROT:Q9PNT0"
 /translation="MEFKLKHKGDMARVCEITTAHSTFLTPFEMPVGTGAVKSLDAN
 DMNEDLAKTILANTYHMYLRPTSKYVGGGLHGFETPRTSGQASLSNG
 SKFPLNGEKSHIDSRHL.FPPEKSYLDQYDPSNDIMLDDIVALLPARKYKISV
 DRILLAKELIITYHKMNQKIGTIGONTREIIOGCTDYERKCAALSNMPRGGLAI
 GGLSVBENALMTEVQNLNPLVDENRPRLMGVGPEDVVENVERVDNEDCMWPR
 NANGNFTFSFGKNTKKAFFINDHEVIDSTGCYCGRNSRQYLHFLAKELTFPR
 LASLHMLHYLEIARAKRREAILNNSTFOGRNPFYH.RGK"
 2596..3363
 /gene="Cj1011"
 2596..3363
 /gene="Cj1011"
 /note="Cj1011, probable membrane protein, len: 255 aa; no
 Hp match"
 /codon_start=1
 /transl_table=1
 /product="putative membrane protein"
 /protein_id="CA873267.1"
 /db_xref="GI:6968447"
 /db_xref="GOA:Q9PNS0"

Db 15765 CTCATTAGCATTTCAAAAAGCCCTATTATTAGTTCATAGATCAGTAA-----AGC 15818
 QY 3199 ACTTTCCTCAGTACCAAAATATTTAACAGATATTTAGTGGTACAGATAGATA 3258
 Db 15819 ACCCTTTTCAAAACCAAAAAGCTCATTCAATCAAAATTTGACAGAAAGGCTGCATAT 15878
 QY 3259 CATAGATTAATTAAGAGTATCATACCCCTGGATTTTATGATTTAGCTAGCAACTAC 3318
 Db 15879 AATAGCAATTAAGAGGCTTATAGAGCGTTTGAATGGGTGAATATACAGAAATATAC 15938
 QY 3315 TTCTTTTCCAACTGCGCAACCAATATAGATATACATCCATGATAGTACTGCTT 3378
 Db 15939 CTCTTTTCTTCTCCTCCTTCTCAAAAACCTTACAGAGCATCAAGTTTAGCGCTT 15998
 QY 3379 AATTAATCATCTTCTCAATTTACATTAACAGAGGAGTTACGACTATTTATATCCTC 3438
 Db 15999 TTGACTTAATTAATTTATTTTCTTAATTAAGAGATTTGAAAAAATCTTTATTTTC 16058
 QY 3439 AAATGCTATTTCAATCTCATCTATTTCTTACGATATCTACAGACTCTA-----TAGC 3492
 Db 16059 TCTTTTCTTCTATTTTCTTTTCTTTTATCGGCTGTTTCTCTTATGATTTTAGC 16118
 QY 3493 TCTCTTACACTATCTTTATCTCTCTCTGTAAGAGCTTTCTATATATATATAGC 3552
 Db 16119 AGCTTAATAGCTTGACTTAAGATATCAACATCAAAAGGCTTGTTAAATAATCTTTAC 16178
 QY 3553 ACCCATATGACAGACTTATACAGAGTGCATATTTACATGCCCATATCATATATAC 3612
 Db 16179 TCCCAACGAAACGCTTCTATAGCAAGGCTTAAGTACATTTCTCTCATATATATATA 16238
 QY 3613 AGGCAATATAGAGTACCTTTCTTAAAGCTTTCCAGTACTTATCCATCATATATGCA 3672
 Db 16239 ATCATATTTATTTTCAAGGCTTTAACAATTTATGCGCATCAATTCCTGCGCATATAT 16298
 QY 3673 TCCCTTAACCATATATATC 3690
 Db 16299 ATCTGAATATCAATATC 16316
 RESULT 33
 AR377574/c 1425 bp DNA linear PAT 18-DEC-2003
 LOCUS AR377574
 DEFINITION Sequence 2580 from patent US 6605709.
 ACCESSION AR377574
 VERSION AR377574.1 GI:40080756
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1425)
 AUTHORS Breton, G.L.
 TITLE Nucleic acid and amino acid sequences relating to Proteus mirabilis
 JOURNAL Patent: US 6605709-A 2580 12-AUG-2003;
 FEATURES
 source
 1..1425
 /organism="Unknown"
 /mol_type="genomic DNA"
 ORIGIN
 Query Match 3.6%; Score 170.6; DB 6; Length 1425;
 Best Local Similarity 50.1%; Pred. No. 1.5e-12;
 Matches 484; Conservative 0; Mismatches 474; Indels 9; Gaps 2;
 QY 2708 ATCATCTCCTTGAAGATTTCAATTAATTCATCTATTAACATTAAGTATGCTT 2767
 Db 1136 ATTAATCTCTTGGCTAGCGCTATTAACGTTAACACAGACATACCTTTTCTAATTTCTG 1077
 QY 2768 AAGTAACTGCGCATTCATATGATGATGCTATTAAGCTTATCATCATTAATATATGA 2827
 Db 1076 AGGTAACTGACCAAGTGTATTCATATGATGATTAATCTTGTGTGTGAGACATTTG 1017

QY 2828 GTACAAAACCTATTTTATACAGATGCTATTCATTAATTAACCTACACATTCGGTATA 2887
 Db 1016 GTTCAACCAAGCTCTGAGCGGTTTTTGGAGAAATATACGGCTAAGCTGGAAAT 957
 QY 2888 TCTGTAATTAATTTCTAATAAGGCTACTTAATTTGAAGACATTTATCTATATAT 2947
 Db 956 TCTTCTGTTCTATCAGCAGAGTGGGGCAATTAATATGATTAATGATGATGATG 897
 QY 2948 AAATCTCAAAAACCTACAGCTTTTACTTCACTTCAATATCTTTGGAAGAAGACA 3007
 Db 896 AGATCTTCAAGAAATACACTTCCATTTAGCGTTTTTCTAATCTTGAAGATCCGGCA 837
 QY 3008 ATATTTCAATCTATATCTAACAAGAACTTATCTATTTTCCCTATATATTTTCCCTCC 3067
 Db 836 ATATAGCAGACATCGACTTTCACAGGGGATACCCCAACGATTAATATATGTCCTTCA 777
 QY 3068 TGTATATCTGAGTATCTTAATTTGCGATAGTATCGTAATCTGTTACTCATCTATA 3127
 Db 776 GCCAAGACGGTATATAGCGGTTGATATCAAGTGCATATCGCAATTTCAATCAGAA 717
 QY 3128 AATTAAGTACCAATGATTTGCTCTATATATTCGAATATGAGAGTACTGTATGAC 3187
 Db 716 AAGAGGAAACCGCATTTGCTTGTCAAAAACGCGCTTGCGTACTGATGATGACACAGTA 657
 QY 3188 AATATATATATCTTCTCCTCAGTACCAATATATTTACCAATATATATAGCTGTAC 3247
 Db 656 AATGCCCTTTTCAAGACCAAAAAGTTCTGATTCATTTAATCTTTGGAAATAGCGCC 597
 QY 3248 ATAGATGATGATATATTAATGAGATATATACCCCTGATTTTATGATATAGC 3307
 Db 596 ATA-----TTGATGCAATTAATAAGGGGCTTGCGCTCGCACTATGCGAATGAAAGCA 543
 QY 3308 CTAGCACTACTTCTTTTCAACACCTGCGCAACCAATATGATATGATATGATGATG 3367
 Db 542 TGTGCAACTTAATTTCTTACAGATCCGACCTGCAATGATATAGACACTAATATGATG 483
 QY 3368 GTAGCTGCTTATTAATCAATCTTCAATTAATGATATGAGAGTATGATGATGAT 3427
 Db 482 CGAAGAAAGCGGCAATATGCGATTAACCTCTTGATGCAAGGGCTTGCGCAATG 423
 QY 3428 TCATA---ATCCTCAATGCTATTTCAATCATATCTATCTTCTATGATATCTACGAC 3484
 Db 422 TCAGACACAGATTAAGGTTAGTTCTGTGATTAAGTTGCTGCTGATAGTGC 363
 QY 3485 TCTATGCTCTTCTTCACTAATTTATCTTCTCTGTAAGGCTTTCTATATA 3544
 Db 362 GTATAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 303
 QY 3545 TCATAGCAACCATATGACAGACTTTACAGAGTGCATATATACATGCCACTATATC 3604
 Db 302 TCAATGCCCTTGCTGATTAAGATTTACCGGGCATCAAGATCAAGATGCCCTGATG 243
 QY 3605 ATATTAAGGCAAAATTAAGATACCTTCTTAAAGCTTTCCAGATCACTTAATCATCA 3664
 Db 242 ATATTAACGGCAACGTAGGTTGCTCTTTAAGCTTTTAAATGAGATTAACCATCA 183
 QY 3665 ATATCAG 3671
 Db 182 ATATCAG 176
 RESULT 34
 AE007680 10861 bp DNA linear BCT 27-JUL-2001
 LOCUS AE007680
 DEFINITION Clostridium acetobutylicum ATCC824 section 168 of 356 of the
 complete genome.
 ACCESSION AE007680 AE001437
 VERSION AE007680.1 GI:15024666
 KEYWORDS
 SOURCE Clostridium acetobutylicum
 ORGANISM Clostridium acetobutylicum
 Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 Clostridium.

REFERENCE		1 (bases 1 to 10861)
AUTHORS	Nolling,J., Britton,G., Omelchenko,M.V., Markarova,K.S., Zeng,Q., Gibson,R., Lee,H.M., Dubois,J., Qiu,D., Hitti,U., Wolf,T.I., Tatusov,R.L., Sabathe,F., Doucette-Scamm,L., Souaille,P., Daly,M.J., Bennett,G.N., Koonin,E.V. and Smith,D.R.	
TITLE	Genome sequence and comparative analysis of the solvent-producing bacterium Clostridium acetobutylicum	
JOURNAL	J. Bacteriol. 183 (16), 4823-4838 (2001)	
MEDLINE	21359325	
PUBMED	11465286	
REFERENCE	2 (bases 1 to 10861)	
AUTHORS	Childress,D., Zeng,Q. and Smith,D.R.	
TITLE	Submitted (24-JUL-2001) GTC Sequencing Center Production, Finishing, and Bioinformatics teams, Genome Therapeutics Corp., 100 Beaver Street, Waltham, MA 02453-8443, USA	
JOURNAL	Location/Qualifiers	
FEATURES	source	1..10861 /organism="Clostridium acetobutylicum" /mol_type="genomic DNA" /strain="ATCC 824" /db_xref="ATCC:824" /db_xref="taxon:1488" 96..476 /gene="CACI702" 96..476 /gene="CACI702" /codon_start=1 /transl_table=1 /product="Hypothetical protein" /protein_id="AAK79668.1" /db_xref="GI:15024667" /translation="MNFTLMKKKKKKRKSVAIVRIPIVTLIIISMAADGIIYYEA YNIIYIKINSKNMVSSEIYNPFKVLVLQNEAKNISDSYLKRITVTNNKISTDEM SMLNKTKGSDSKVEYNIFPAQL"
gene	473..823 /gene="CACI703"	
CDS	473..823 /gene="CACI703" /codon_start=1 /transl_table=1 /product="Methyl-accepting chemotaxis protein (fragment)" /protein_id="AAK79669.1" /db_xref="GI:15024668" /translation="MTSPVITFPANPIDDGNLTGVGKTIPDYFSKRPDSFKWNG NSTKLISLNAALBASRLBEGKFGVASEIKKLSNNVETIYVIGIAVQAINEKYVN MKDKTSLNRDXYD"	
gene	824..1039 /gene="CACI704"	
CDS	824..1039 /gene="CACI704" /codon_start=1 /transl_table=1 /product="Hypochemical protein" /protein_id="AAK79670.1" /db_xref="GI:15024669" /translation="MAAYNTKTNVNFPSLIQELINNKKIVDENANVSSQNKIYH EKNISAFIDERNSTIELEBNIGFRLL"	
gene	1188..2084 /gene="CACI705"	
CDS	1188..2084 /gene="CACI705" /codon_start=1 /transl_table=1 /product="Periplasmic phosphate-binding protein" /protein_id="AAK79671.1" /db_xref="GI:15024670" /translation="MKSKKIKLVVAITMTIAGLPVCGNSNSSKSSSTTTKAYS GSTTSAGTALPLAEKSYDPKVSVVSGSGTGLOVLAIVEVGSDF AEERLNADDKALTDRKVCALGEAVTNNDVVNLTKQVODIFGTIRNMNVGGP SEAINIVHPKSSGTRATPTTKTVMGCKKDQIGTQDASSVTKAMTSKSGSIYA FSLVTDDEKDIINISLDGVEATISKVNBNGTYTPWSTEHHTYTGKGNDAKAPIDYM MSSNDALLERKGYSASNTKK"	
gene	2183..3073 /gene="CACI706"	
CDS	2183..3073 /gene="CACI706" /codon_start=1 /transl_table=1 /product="phosphate permease" /protein_id="AAK79672.1" /db_xref="GI:15024671" /translation="MEKRSFLKUKTEYVGGRFATICGPVITLTIISPIASKGIS IFTKHYSIFDLVSTTWPKDSKTPKGLFILEGSTWSVIGAVITSPVALAFAV NIIISPGLDVRILKPSLELPVGIPSVVGWGLVLPVPINKRGVGFSAAILAFV IMILEPTIASLSDAIKVMENYEIGSYGLGATRMOIQIIXIIISSKNGILTGLGA RAFGALAOMWGIGVYSAGQLFSPPTLITSVLTMANTPNGTAWNDALMSLAFLL LVISRFIVIRIFGRKEV"	
gene	3073..3957 /gene="CACI707"	
CDS	3073..3957 /gene="CACI707" /codon_start=1 /transl_table=1 /product="permease component of ATP-dependent phosphate uptake system" /protein_id="AAK79673.1" /db_xref="GI:15024672" /translation="MKSKLYDKIATIIILYLSLFVILLAAFIIFIIRGRDSLNSF LFNPKIAREKGGIGPELIPVYMLISLITVIGIAGIYMSYAKERINLFIPL CLETMASPLSIVGLFGLVFTWMGWGITLSCALVITLINPSMRVSENARASS SKVVASIGLGATWTMOTIKNIVIPSNBEITGTIISAGRIFGBAALITAGMAPN INFDTISVDYKTSAPSILPRPATITLAHIWKLNBSGMIPDKTIANGSSAVLVMTLFF NMFAIRIKRTKYSLSGK"	
gene	3959..4718 /gene="CACI708"	
CDS	3959..4718 /gene="CACI708" /codon_start=1 /transl_table=1 /product="ATPase component of ABC-type phosphate transport system" /protein_id="AAK79674.1" /db_xref="GI:15024673" /translation="MGIIOTKDLNYGYNOALKINLDEPSANTVTALLISPGCSGKST FLRTNRMNDLISVTKIDGEMFEKGVYDVDEIELRKVGWFQKPNFPMISYDN VAYGRIRIGIKRKKDBEIVRSIKSALMEYVDRKSKALGSGGCCOCLTARTL AVEEPVLMDPEPSALDPISTLKTEBLMDLKKKITVIIVTHNQAGRISDNTAPFL NGEVVENKGTEDIYPKPDKTEDYITGRFG"	
gene	4750..5409 /gene="CACI709"	
CDS	4750..5409 /note="phcy" /gene="CACI709" /codon_start=1 /transl_table=1 /product="phosphate uptake regulator" /protein_id="AAK79675.1" /db_xref="GI:15024674" /translation="MTKRIPSDDLSEHLHSLLRWGSAEKQIYDCMEALEKODENNAE VIHKDDIIDMQKEIENKIRILRNMQOPVABEILNIPITVKIYVLDLEIGHANDI AKAIRLNGEGRHDIVKEIYMMGNKVSMTKDSQSLVERLRDRAIVCRRDDVDSL YKRIFNELNMTMSDEKSRVQLTFPCVYLERIGRTTNVCSITIIYLTGKQVDN D"	
gene	5586..6899 /gene="CACI710"	
CDS	5586..6899 /gene="CACI710" /codon_start=1 /transl_table=1 /product="Fe-S oxidoreductase, related to NifB/MoaA family with PZD-N-terminal domain" /protein_id="AAK79676.1" /db_xref="GI:15024675" /translation="MENKIYKIEPOSJAEMGIEVDSILISINGEKVIDIDRFMA DEYVAVIKQTOKMEIWEIJEKYEIDKGLVFEBGIIIDSAKSCRNKIIFCFIDQLKRG	


```

/mol_type="genomic DNA"
/strain="297"
/db_xref="taxon:139"
223..1110
/gene="nlpd"
223..1110
/gene="nlpd"
/feature="nlpd"
/feature="putative membrane-bound metalloproteinase; similar
to ORF B80761 in Borrelia burgdorferi strain B31 genome"
/codon_start=1
/translation_table=11
/product="NLPD"
/protein_id="AA021217.1"
/db_xref="GI:33520758"
/translation="MSKIFLEFVAGFPEFLKIIYVSPYSEIKNSQDPVPSDLKIKVL
KNNKQHIPLFEPYVYKGGKDDPEFKANKINGOSGATINILDSPAVSQGEIYPS
KKGVPEDSKDYRKNLLATRDLTAEKVKIKENDRYVEFEFDPKKNPDSFGST
ELIFLPIANRIPFLKKEIVISDRGPNDFGKNSPHNGIDLAAPMAEVSSSGIV
IEKYNIDYGNFVYVGHKNNIKSLYGHMLYVYKIGDPKSEBFGXVQGTGRATGPH
LHFELIKNIPINPLKFLK"
/complement(1359..2717)
/gene="xrp2"
/complement(1359..2717)
/feature="xrp2"
/feature="response regulator; enhancer-binding protein that
activates sigma 54-dependent genes; transcriptional
activator that controls expression of sigma 54"
/codon_start=1
/translation_table=11
/product="Rip2"
/protein_id="AA021218.1"
/db_xref="GI:33520759"
/translation="MSKIVADDEKNIKRECIATYLEDSEYFVPTASDGEALETIENE
NLDTVILSDRMPQISEKILKIVKXNLIPIILATVVSADVAMEGVEFLTK
PLDLERLLIKRLKKNKNDNDENLILRIKDLKYEIMKSLMOKYFELV
KIAKNSAVLITGESVGEKILADAFEDLSNRDQFLKVAALASELSEELGHE
KAGFASISKKRPELANKGTIFDELDEISPEIQVLRVLOKNTFSEVSGEATIK
VDRLAATKNIEEIKKEKPEDEIYFRLININIPERKNDISLNTILKDV
AKENNEEKTLSNDAMKALYYDMRGNIIEIKNVLASALISKQGTREDEIPAKIK
NENLIRKIDFISLSEAKELIKQTLPHSKNKSCKAEILKIGKRLANKITEYNI
DO"

```

Query Match 3.64; Score 169; DB 1; Length 3900;
 Best Local Similarity 49.54; Pred. No. 1.7e-12;
 Matches 600; Conservative 0; Mismatches 575; Indels 36; Gaps 5;

ORIGIN

```

1886 CAATTATATGAGCTTCGCCCAACAGCTTCAAAAGTTTGTTCGAGTACTCTTA 1945
3081 GTAATCTTAATTTGGATGATGATGATGATGATGATGATGATGATGATGATGAT 3140
1946 AAGTTGACTTGAATTTACAGTCAAAATTTCTGTAATCTATCAAGAAAATTTGCTT 2005
3141 GATTGCTTGTCTTAATTTCAATATGAGAGGAGTACTATGAGACATATATTTATAC 3200
2006 TGTTCGCAAGTTCAAAATCTGCTTTTGTTCGAAAATGCTCCAGTAATGCT-----C 2059
3201 TTTCCTCACTCCAAATATATTAACAGATATTTAGCTGGAGATGATGATGATGATGAT 3260
2060 CTTTTCATGCGCAAAAGTTCACCTTCATGAAATGCTTTCAGAAAGTCTCGCAATTTA 2119
3261 TAGATATTAATGAGATATCAATACCTTCGATTTTATGATATGAGCTGCAACTACT 3320
2120 CTTTATATTAATGATTTGATCTTATTTGAAAGATCAAAATATGATCTGCTATTTT 2179
3321 CTTTCCACACCTCGGCAACAGATATGATATGATGATGATGATGATGATGATGATGAT 3380
2180 CTTTACCAACACCTTCGCTGCTGTTTAAAGAACATGATGATGATGATGATGATGAT 2239
3381 TAATCATATCTTCTCAAAATATGATATGATGATGATGATGATGATGATGATGATGAT 3437
2240 TTACAGATGATCAAAATTTTTCATATATGATGATGATGATGATGATGATGATGAT 2299
3438 CAAATGCTGATTTCACTCATATTTTCTGATGATGATGATGATGATGATGATGATGAT 3486
2300 TTAAATCTTTCTTATATGATATATTTCTAAATGATGATGATGATGATGATGATGAT 2359
3487 -----TATGCTCTCTTACAACTTATATCTTCTGTAAGGCTTTCTA 3539
2360 TTTATTTATGATATCTTTTATTTATTTGCAAAAGCTTTTCAAGCTCTAAGGCTTTGTA 2419
3540 TATATATCAATGACCAATATGACAGATTTTACAGAGTGGCAATATTCATGATGATGAT 3599
2420 AAAATATCAATGACCACTCTCTGATGATGATGATGATGATGATGATGATGATGAT 2479
3600 TATATATCAATGACCAATATGATGATGATGATGATGATGATGATGATGATGATGAT 3659
2480 TTAAATATATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2539
3660 CATCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3719
2540 CAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2593
3720 AAGCATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3779
2594 CAATGTTTCAAGAGCTTCTTCGCTGATGATGATGATGATGATGATGATGATGATGAT 2653
3780 TTAATATATCTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3839
2654 CAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2713
3840 AAATATCAAT 3850
2714 TCATTTTCAAT 2724

```

RESULT 36
 AC120883 133877 bp DNA linear HTG 06-AUG-2002
 LOCUS Homo sapiens chromosome 18 clone CTD-2146H24 map 18, WORKING DRAFT
 DEFINITION
 ACCESSION AC120883
 VERSION AC120883.2 GI:22123189
 KEYWORDS HTG; HTGS PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 133877)
 Birren,B., Nusbaum,C. and Lander,E.

TITLE Homo sapiens chromosome 18, clone CTD-2146H24
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 133877)
AUTHORS Birren,B., Linton,J., Nusbaum,C., Lander,E., All,A., Allen,N.,

TITLE Direct Submission
JOURNAL Submitted (13-MAY-2002) Whitehead Institute/MIT Center for Genome
REFERENCE Submitted (13-MAY-2002) Whitehead Institute/MIT Center for Genome
AUTHORS Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 133877)
Birren,B., Nusbaum,C., Lander,E., All,A., Allen,N., Anderson,S.,

TITLE Direct Submission
JOURNAL Submitted (06-AUG-2002) Whitehead Institute/MIT Center for Genome
COMMENT On Aug 6, 2002 this sequence version replaced 91:20531967.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIRB
Web site: http://www-seq.wi.mit.edu
Contact: sequence submissions@genome.wi.mit.edu
Project Information
Center project name: 126768
Center clone name: 2146_H24
Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 13264 bases at least Q40
Consensus quality: 13263 bases at least Q30
Consensus quality: 13305 bases at least Q20
Insert size: 13300; agarose-ef
Insert size: 13357; sun-of-ctnigs
Quality coverage: 17.1 in Q20 bases; agarose-fp
Quality coverage: 17.0 in Q20 bases; sun-of-ctnigs

NOTE: This is a 'working draft' sequence. It currently
consists of 4 contigs. Gaps between the contigs
are represented as runs of N. The order of the pieces
is believed to be correct as given, however the sizes
of the gaps between them are based on estimates that have
provided by the submitter.
This sequence will be replaced
by the finished sequence as soon as it is available and
the accession number will be preserved.
1
717 816: contig of 716 bp in length
817 1487: contig of 671 bp in length
1488 1587: gap of 100 bp
1588 1993: contig of 406 bp in length
1994 2093: gap of 100 bp
2094 133877: contig of 131784 bp in length.
Location/Qualifiers
1. 133877
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="18"
/map="18"
/clone="CTD-2146H24"
/clone_11b="CTD1 Human BAC"
1. 716
/note="assembly_fragment"
817. 1487
/note="assembly_fragment"
1588. 1993
/note="assembly_fragment"
2094. 133877
/note="assembly_fragment"
ORIGIN
Query Match 3.6%; Score 168.8; DB 2; Length 133877;
Best Local Similarity 45.3%; Pred. No. 5e-13;
Matches 983; Conservative 0; Mismatches 1152; Indels 33; Gaps 9;
QY 322 ATCCATTTTCCAAATTTTACAAATCATGAGTTCTTTATTCACATCATTTAA 381
Db 94004 ATGAT 94063
QY 382 ATCCATCAAAACAAAT 441
Db 94064 ATAAAAAT 94123
QY 442 GCAAGAACTCAAAAT 501
Db 94124 ATAAAT 94183
QY 502 TACCATCATATTTAGAT 561
Db 94184 TAAAT 94242
QY 562 TAACTAT 621
Db 94243 AT 94302
QY 622 TAGTAT 681
Db 94303 AT 94362
QY 682 TGCACTGTGACAT 738
Db 94363 GAT 94422
QY 739 TACCAT 798
Db 94423 TTTAT 94482
QY 799 CAACATATATGACACGACATATATATATATATATATATATATATATATATAT 858

[illegible]

Db	195539	ATATTTATTTGATTAATTTATTTATTAATAATAAATAATATATTTATTAATAATAATAATAATAATAA	955580
Qy	1938	TAACTAATTAATAAAGAACCAATACATATATTAATAATTAATAAAGCACTATTATTAATCTAGTAA	1997
Db	955599	TGTAAATATAAATTTAAAGTGA--ATATAATATATATTTATTAATAATAATAATAATAATAATAT	956555
Qy	1998	TATTAACATATGATCTCTGACATATATAATTTCAATCATCTCTGACATATTTCAACATATAA	2057
Db	95656	TTATATATATATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA	957155
Qy	2058	ATATGATAAACTGTTTTTTTATATGCCATTAGTCTTAATAAAATTGCTAATTACCCAAA	2117
Db	95716	AATATATAATTAATTAATTTATTTATTAATATATATATTAATAATAATAATAATAATAATAATAA	957755
Qy	2118	AGGTTAAATATACATCTTCTCGACACAATGTGAAGAGAATACATCAAGCTCAATATAAT	2177
Db	95776	ATATATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT	958355
Qy	2178	AAACAAATTTCTTAATATAAACAACCTACTCTAAACATAGAACCAATAGATATCATACCT	2237
Db	95836	ATATATAATATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATTT	958955
Qy	2238	AAGTAAGATATCCCTTCACAGATGTACAAATATAAAGAAAATGCTTATGTATTTAT	2297
Db	95896	ATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT	959555
Qy	2298	ACATATACAAATTTTCCACTTCACACCAACAAACCAAGTCTCAT-TATAACAATTC	2356
Db	95956	TTATATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT	960155
Qy	2357	AGAAATATATATATCTTCACCAACACACTGATTAACAATTTCTATAATPACTATATCTTT	2416
Db	96016	AATATAATTTATTTATTTATTAATAATAATAATAATAATAATAATAATAATAATAATAATAAT	960755
Qy	2417	TTAAAGATATGCTTCACATTCATACATAAATAATTAATATACATGCAATGTTATTCAGAA	2476
Db	96076	ATTTAAATATTAATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTA	961355
Qy	2477	ACATTACA	2484
Db	96136	TATTTATA	96143
RESULT 37			
PFML4P3			
LOCUS	PFML4P3	349751	DNA linear INV 29-JAN-2003
DEFINITION	Plasmodium falciparum ML4P3.		
ACCESSION	AL035476	AL844503	
VERSION	AL035476.7	GI:23498195	
KEYWORDS	Plasmodium falciparum 3D7		
SOURCE	Plasmodium falciparum 3D7		
ORGANISM	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.		
REFERENCE	1		
AUTHORS	Hall, N., Pain, A., Berriman, M., Churche, C., Harris, B., Harris, D., Mungall, K., Bowman, S., Atkin, R., Baker, S., Barron, A., Brooks, K., Bucke, C.O., Burrows, C., Cherevach, I., Chillingworth, C., Chillingworth, T., Christodoulou, Z., Clark, L., Clark, R., Corcoran, C., Cronin, A., Davies, R., Davis, P., Dear, P., Dearden, F., Doggett, J., Felwell, T., Goble, A., Goodhead, I., Gilliam, R., Hamlin, N., Hance, Z., Harper, D., Hauser, H., Hornaby, T., Holroyd, S., Horrocks, P., Humphrey, S., Jangori, K., James, K.D., Johnson, D., Khoronou, A., Knights, A., Konstantinov, B., Kyes, S., Larke, N., Lawson, D., Lennard, N., Line, A., Meddison, M., Mclenn, J., Mooney, P., Moulle, S., Murphy, L., Oliver, K., Omond, D., Price, C., Quail, M.A., Rabinovitch, E., Rajandream, M.A., Rutter, S., Rutherford, K.M., Sanders, M., Simmonds, M., Seeger, K., Sharp, S., Smith, R., Squares, R., Squares, S., Stevens, K., Taylor, K., Tiley, A., Unwin, L., Whitehead, S., Woodward, J., Sulston, J.E., Craig, A., Newbold, C. and Barrell, B.G.		
TITLE	Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13		
JOURNAL	Nature 419 (6906), 527-531 (2002)		
MEDLINE	2225708		

PUBMED 1236867
2 (bases 1 to 349751)
REFERENCE Hamlin, N., Pain, A., Berriman, B., Hall, N., Bowman, S., Churcher, C.,
AUTHORS Harris, B., Harris, D., Lawson, D., Quail, M. and Barrett, B.
TITLE Direct Submission
JOURNAL Submitted (24-FEB-1999) P.falciparum Genome Sequencing Consortium,
The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge
CB10 1SA, UK
COMMENT On Oct 3, 2002 this sequence version replaced gi:5763808.
For more information about this sequence or the Malaria Project,
see http://www.sanger.ac.uk/Projects/P_falciparum.
FEATURES
source
location/Qualifiers
1..349751
/organism="Plasmodium falciparum 3D7"
/mol_type="genomic DNA"
/isolate="3D7"
/db_xref="taxon:36329"
/chromosome="4"
/clone="YALAP3"
/complement(join(2130..3551,4472..9988))
/gene="PF0630c"
/note="synonym: VAR"
/complement(join(2130..3551,4472..9988))
/note="similar to Plasmodium falciparum var. mal4p2.58
mal4p2.58, var SMALL:09U065 (EMBL:AL035475) (2277 aa)
fasta scores: E(): 1.5e-143, 52.77% id in 2486 aa, and to
Plasmodium falciparum variant-specific surface protein
var-7 SMALL:Q26034 (EMBL:U42636) (2182 aa) fasta scores:
E(): 1.9e-126, 46.72% id in 2412 aa"
/codon_start=1
/product="erythrocyte membrane protein 1 (PfEMP1)"
/protein_id="CAD49166.1"
/db_xref="GI:23498197"
/db_xref="GOA:Q811V7"
/cd_xref="SPTREMBL:Q811V7"
/translation="MGPSPTADPVSQSAKHALDSIGKRVHQAQVONAKORSNGDLKGL
TSATLSGGSAPFENPCILIKKDELLIGSRRHPCNGSASERFSEVSGGDDK
KTEGRRNGGACAPYRPLFLCNKMEKGRPTTKHLLDVCNAYEASLIIRH
DOTDATYPSGDSFSCMTMLARSPADIGDITRGDLVGGKKKKKTEEDQLESFKKI
FGDINELTNGRNGVADHODNGGAYFQLEDEWNTANRATVWKAITCAGATGAYFR
PICSNRQPSQAHYCRGNDGKDDKPTDPTFYDVPOLRWFEWMAEFCRKK
IYVGVTKYCRKYSNGEPRYCRSGYDCTKTRAIKRYRMNGCISCLACAPYVD
MINNOKEQDKKKYTDIENASRSRQKRGASTGSSSYDGESEKFNILKDYD
GVYDFDLKLNKESQAVKNDGGTINTEKNDKNNNNKDKGFYSSEYCOVCPDC
GVYNGSGWEKKKXNDCKIKYRKKQAPRPIKILSGKEKEIEKTLAFDCKD
GGSNDSLYDPWQCFDQLEKEKEBEGVDNRYNDRVTSGLCITLQKRGENGKK
QTYNDFNFWAHMLKDSIHWKKLEKLGQITRCRNNKCKECCORVWEKK
KTEGKIKDHFNTONIGETNCDPTVLEGVKQIFLNDSYQDKQNLDEBEKHLK
HSEMLOETSGDLTCGASDNKETIMDLIHEBGIACECLAKNECEKAKEGRS
DSHDPEQPDADNEDLDDEDEDEEVEVDNTEOEGEPVYVQCEGSSPTPA
PAGPVYCDIYVQLEFSPQSFACITLKYTGKNGMKIPGNTSDTTGSSPATGA
GRRDSDGDTGCGKATGCGICIPRRRLIYVGLSOMASRTQGTGSSSGSNILE
ATIQSAIITFFLHMYKAKENTKTGVAGADPTPATSVATAPAPAVSRSLCL
LSEGVGEPGMAVGVKSIPEPPLGVAGGIPGVALGGAGAPGVALGVGQAO
PILRLPGVANGLOSPOSLRLTLGHPAGSGEDTPQWQGHIPDRLRMYET
LADYDICIQGRDITVGTIVNTSGSSSKIKISEKIEILNHKNEKAPKSVET
KPTTEWMSNGGHIWNGMICALTYDSSGISTEDVDKILIDKDTGPOKNGNDY
YERKLLKDESGPKNDITOPATLKDFEILPYPRMLHEMGSFGCAKAPLKVYKH
CENIDKAGHHYSGSGYDCTPRVIERNKRVNLGLCYKQCRKKKIDIKPYEYH
QEKIKDEYGLTKDSSDDKLEGGKCHENPLKELKACRSEDNNDDNKNKINDEK
KTFNPSKYCAKAPYGVKTYGNYIPNEKDKYSKRGVKEKENDTIPNILEVLVGR
GEKEDKHLDAKNTGLFEVARYEQWNCQKKGIIDOCKITKFNANDIDFDDIVFNE
FQPMRLYFVQDYNLKOKINCTKETEKEKEKYSCTOGCNDKECVKMLKCKO
EMTQITLYKQYSKISDQIAPRVKSYFVDQDLPNDYKAOEVEKDEKELMGCT
GNLKEGEGPKGCMGDFITNLISLQKKIDCNQAOANSVETQPSENAAOCODTH
PDDEDLLEENNOVABOINCPNOVEKIEEVEKETAOTTAETRAAAGRROR
PAPAPAPSPRPRLPKRPPRPDLPALNKMSTIMSGYIGRAAPYFPLAKKT
KASVGNLFQILQPKGDYDIPUKSSNRITPASRPGKTYIHWGSSSDERAKFM
SOTVDVLTSESEYELDVNDIVPGSPKTKLIEVLPSGNTPTSPIDNPTPO
PLTDEMNQLEKDFISMLQNTQNTENLHDVNDNTHPTWSRMDQKFLISIDH
RLFSGEENYDMFNSGNPINSIDSTNSMDLTSSNHSYNDKIDLQSGDLINDAL
SGNHIDYDEMCKRENELFGTHHPKITSNRVVTQTSDDP:INQNLPHTWLDRH

misc_feature
gene
CDs
13697..13898
/note="GC-rich region"
/complement(join(17269..18690,19611..25169))
/gene="PF0635c"
/note="synonym: VAR"
/complement(join(17269..18690,19611..25169))
/note="similar to Plasmodium falciparum var. mal4p2.58
mal4p2.58, var SMALL:09U065 (EMBL:AL035475) (2277 aa)
fasta scores: E(): 1.1e-129, 49.19% id in 2498 aa, and to
Plasmodium falciparum variant-specific surface protein
var-7 SMALL:Q26034 (EMBL:U42636) (2182 aa) fasta scores:
E(): 7.9e-126, 46.25% id in 2441 aa"
/codon_start=1
/product="erythrocyte membrane protein 1 (PfEMP1)"
/protein_id="CAD49167.1"
/db_xref="GI:23498197"
/db_xref="GOA:Q811V6"
/cd_xref="SPTREMBL:Q811V6"
/translation="MAPTSGGGETKDSAKHALDSIGKRVHQAQVONAKORSNGDLKGL
LTSATLSGGSAPFENPCILIKKDELLIGSRRHPCNGSASERFSEVSGGDDK
DKTEGRRNGGACAPYRPLFLCNKMEKGRPTTKHLLDVCNAYEASLIIRH
DOTDATYPSGDSFSCMTMLARSPADIGDITRGDLVGGKKKKKTEEDQLESFKKI
FGDINELTNGRNGVADHODNGGAYFQLEDEWNTANRATVWKAITCAGATGAYFR
PICSNRQPSQAHYCRGNDGKDDKPTDPTFYDVPOLRWFEWMAEFCRKK
IYVGVTKYCRKYSNGEPRYCRSGYDCTKTRAIKRYRMNGCISCLACAPYVD
MINNOKEQDKKKYTDIENASRSRQKRGASTGSSSYDGESEKFNILKDYD
GVYDFDLKLNKESQAVKNDGGTINTEKNDKNNNNKDKGFYSSEYCOVCPDC
GVYNGSGWEKKKXNDCKIKYRKKQAPRPIKILSGKEKEIEKTLAFDCKD
GGSNDSLYDPWQCFDQLEKEKEBEGVDNRYNDRVTSGLCITLQKRGENGKK
QTYNDFNFWAHMLKDSIHWKKLEKLGQITRCRNNKCKECCORVWEKK
KTEGKIKDHFNTONIGETNCDPTVLEGVKQIFLNDSYQDKQNLDEBEKHLK
HSEMLOETSGDLTCGASDNKETIMDLIHEBGIACECLAKNECEKAKEGRS
DSHDPEQPDADNEDLDDEDEDEEVEVDNTEOEGEPVYVQCEGSSPTPA
PAGPVYCDIYVQLEFSPQSFACITLKYTGKNGMKIPGNTSDTTGSSPATGA
GRRDSDGDTGCGKATGCGICIPRRRLIYVGLSOMASRTQGTGSSSGSNILE
ATIQSAIITFFLHMYKAKENTKTGVAGADPTPATSVATAPAPAVSRSLCL
LSEGVGEPGMAVGVKSIPEPPLGVAGGIPGVALGGAGAPGVALGVGQAO
PILRLPGVANGLOSPOSLRLTLGHPAGSGEDTPQWQGHIPDRLRMYET
LADYDICIQGRDITVGTIVNTSGSSSKIKISEKIEILNHKNEKAPKSVET
KPTTEWMSNGGHIWNGMICALTYDSSGISTEDVDKILIDKDTGPOKNGNDY
YERKLLKDESGPKNDITOPATLKDFEILPYPRMLHEMGSFGCAKAPLKVYKH
CENIDKAGHHYSGSGYDCTPRVIERNKRVNLGLCYKQCRKKKIDIKPYEYH
QEKIKDEYGLTKDSSDDKLEGGKCHENPLKELKACRSEDNNDDNKNKINDEK
KTFNPSKYCAKAPYGVKTYGNYIPNEKDKYSKRGVKEKENDTIPNILEVLVGR
GEKEDKHLDAKNTGLFEVARYEQWNCQKKGIIDOCKITKFNANDIDFDDIVFNE
FQPMRLYFVQDYNLKOKINCTKETEKEKEKYSCTOGCNDKECVKMLKCKO
EMTQITLYKQYSKISDQIAPRVKSYFVDQDLPNDYKAOEVEKDEKELMGCT
GNLKEGEGPKGCMGDFITNLISLQKKIDCNQAOANSVETQPSENAAOCODTH
PDDEDLLEENNOVABOINCPNOVEKIEEVEKETAOTTAETRAAAGRROR
PAPAPAPSPRPRLPKRPPRPDLPALNKMSTIMSGYIGRAAPYFPLAKKT
KASVGNLFQILQPKGDYDIPUKSSNRITPASRPGKTYIHWGSSSDERAKFM
SOTVDVLTSESEYELDVNDIVPGSPKTKLIEVLPSGNTPTSPIDNPTPO
PLTDEMNQLEKDFISMLQNTQNTENLHDVNDNTHPTWSRMDQKFLISIDH
RLFSGEENYDMFNSGNPINSIDSTNSMDLTSSNHSYNDKIDLQSGDLINDAL
SGNHIDYDEMCKRENELFGTHHPKITSNRVVTQTSDDP:INQNLPHTWLDRH

QY 1980 CACCTATTACTAGTAATTTACATAGTCTCCGACATATAATCTCATCTCTG 2039
 Db 68746 AATTATTGAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTT 68805
 QY 2040 ACATATTTCAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2099
 Db 68806 AATATATTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 68865
 QY 2100 ATTGCTAATTTACCAAAAAGTTAAATTCACATTCCTGCAACAAATTAAGAGATCA 2159
 Db 68866 AATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 68925
 QY 2160 ATACAGCTCAAAATTAATTAACAAATTTCTAATATACACCTACTCAAAATTAAGAAC 2219
 Db 68926 AT-TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTG 68984
 QY 2220 CAATAGTATATCATACCTAAGTATCTCTACAGCATGTATCAAAATTAATTAAGAAA 2279
 Db 68985 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 69044
 QY 2280 ATCGTATGTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2339
 Db 69045 TAATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 69104
 QY 2340 CTATTAATTAACAAATTCAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2398
 Db 69105 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 69164
 QY 2399 TAATTAATCATATCTTTTAAAGAGTACTTCAATTCATCATTAATTAATTAATTA 2458
 Db 69165 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 69224
 QY 2459 TCGAATGTATTAATTAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2510
 Db 69225 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 69284
 QY 2511 -----GTAAATGCTGACGCTTCATTCACAAATTCAGAGTGTGATACATTAATTC 2564
 Db 69285 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 69344
 QY 2565 CAAAACGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2624
 Db 69345 TTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 69404
 QY 2625 ATGGTAACAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2684
 Db 69405 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 69464
 QY 2685 CTACGTGTAATCTTTTGTGATATCATCTCCCTGAGAGATTCATTAATTAATTAATTA 2744
 Db 69465 ATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 69524
 QY 2745 CTATTAACATTAAGTAATGTCTTAAGTACCTGGCCATTCATTAATTAATTAATTA 2804
 Db 69525 TGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 69584
 QY 2805 AAGCTCATCATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2858
 Db 69585 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 69644
 QY 2859 TCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2918
 Db 69645 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 69704
 QY 2919 TAATTAAGAGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2978
 Db 69705 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 69764
 QY 2979 CACTTAATTAATCTTTGAGAGAGACAAATTAATTAATTAATTAATTAATTAATTAATTA 3038
 Db 69765 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 69824

QY 3039 TACTATTTTCCCTATATTTTCCCTCGTAATATCGAGTAATCTTAATTTGCGTAT 3098
 Db 69825 TATATATTAATTAATTTTATTTTATGATACAAATCTTAATACACACTTAATTAATTAAT 69884
 QY 3099 CGTATGTAATCTGTACTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3158
 Db 69885 ATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 69943
 QY 3159 TTCCAAATTAAGAGGATCTCTATGACAAATTAATTAATTAATTAATTAATTAATTAAT 3218
 Db 69944 TACATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 70003
 QY 3219 TATTAACCAAGTAAT 3234
 Db 70004 TTTATTAATTAATTAAT 70019

RESULT 38
 AX622666
 LOCUS AX622666 996 bp DNA linear PAT 20-FEB-2003
 DEFINITION Sequence 5629 from Patent WO02094868.
 ACCESSION AX622666
 VERSION AX622666.1 GI:28450652
 KEYWORDS
 SOURCE Staphylococcus aureus
 ORGANISM Staphylococcus aureus
 Bacteria; Firmicutes; Bacillales; Staphylococcu.

REFERENCE
 1. Mesigian, V.C., Mora, M.C. and Searse, J.M.C.
 Staphylococcus aureus proteins and nucleic acids
 Patent: WO 02094868-A 5629 28-NOV-2002;
 Chiron Spa (IT)

FEATURES
 Location/Qualifiers
 source 1..996
 /organism="Staphylococcus aureus"
 /mol_type="unassigned DNA"
 /db_xref="taxon:1280"

ORIGIN
 Query Match 3.5%; Score 168.6; DB 6; Length 996;
 Best Local Similarity 51.4%; Pred. No. 3.1e-12;
 Matches 505; Conservative 0; Mismatches 454; Indels 24; Gaps 4;

QY 635 AATATCGATTTTATGAGTGCAGATCATTTGGCAGCAATAGCAATTCAGTGCAGCA 694
 Db 7 AATATTCGCTTTTGTGATAGGAAATTTGGACAGCCCTTCCATATGTTCTTGAGAA 66
 QY 695 CATGTATATCACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 754
 Db 67 AATGATCATGATGTTTGTGATGGGTAATTAATTAATTAATTAATTAATTAATTAATTA 126
 QY 755 TACGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 814
 Db 127 TGTATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 186
 QY 815 AGCAATTAAGCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 874
 Db 187 TCAGATATGACCAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 243
 QY 875 CAAATTAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 934
 Db 244 AAGCAATTAAGCAAGTGTCTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 297
 QY 935 CCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 994
 Db 298 ACTTTATACATGTTGCTTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 357
 QY 995 GCGAAGAAATTT-----ACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1045
 Db 358 ATGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 417
 QY 1046 TTGCTTAAGAAATTAAGCAATCTTCTGTGATAGTACTCTGCTGATTAATTAATTA 1105

Db 418 CATGGGAGAGAGTTGAGTCAGCAACCACTACAGTCTGCTTCATCAAAAGATATA 477

Qy 1106 GAACCTGGTGAATCATGATAGAAACATAGATGATGCTTCAAAAATATATACCAT 1165

Db 478 AGCTAGTAAATTAACGCAAGATTTATATGATATATTTGGGTGTACAGAAAT 537

Qy 1166 CAAGATTTATAGTGTACAGATTTGAGCTGCATTTAAGAACATTAATGCAATTGATGT 1225

Db 538 GATACCTGATTTGTTGTAAGCTGTGTGATGCAATTAATCAACGAGTAGCAAGT 597

Qy 1226 GGAATTAATCGCTGCAAAAAATTTAGTAATATATCTGCTGCTTATATACAAAGC 1285

Db 598 GGTATGCTGATGCAATTTGGCTAGCTGATATATCAAAAGCTGATTAATGACTCGGCG 657

Qy 1286 ATGAATGAATTAACACATATATATAGCAAAAAATATCATATAGATCTTCAATCTTA 1345

Db 658 TTACGGGAAATTAAGTGA-----TTAGTGAAAGATTAAGTGGCGATCCTATGACATTT 711

Qy 1346 ATGGTGCATGATGCTTGAAGATCTATATATTAATATATCAACAGAACATTCAGCAAT 1405

Db 712 CTAGGTTTATAGTGTATGCTGATGCTATATCTTCTTGCATATACATATCTCGGAAT 771

Qy 1406 ATGGCTTTGACATGAATATAGAAAAGTGAATATTAATATATATATATATATATATAT 1465

Db 772 TTACATATAGATATTAATCTGCAAGTGAATCAATGATCAAGCATTAATCTGAATG 831

Qy 1466 CTAAAGCTTTGAGAGAACACGATCTGTAATACCACTGATATATATATATATATATAT 1525

Db 832 AATATGTTGTTGAGGATTTATATCAACATTAATGATTTATGCTTATGCTTATGAGAAAA 891

Qy 1526 AATGAGAACATCAATTTGATATCTATATATATATATATATATATATATATATATATAT 1585

Db 892 AATGAGATGATGCCAATTAACAAAGCATTAATATATATATATATATATATATATATAT 951

Qy 1586 GATAAGCATATATCAACATATTT 1608

Db 952 AAGAGATGCTGAAGATTTAAT 974

RESULT 39

PFMALIP3 67970 bp DNA linear INV 29-JAN-2003

LOCUS PFMALIP3

DEFINITION Plasmodium falciparum DNA from MALLIP3.

ACCESSION AL031746 AL844501

VERSION AL031746.9 GI:6594243

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1. Plasmodium falciparum 3D7

Plasmodium falciparum 3D7

Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.

Hall, N., Pain, A., Barriman, M., Churcher, C., Harris, B., Harris, D., Mungall, K., Bowman, S., Atkin, R., Baker, S., Barron, A., Brooks, K., Bucee, C.O., Burrows, C., Cherevach, I., Chillingworth, C., Chillingworth, T., Christodoulou, Z., Clark, L., Clark, R., Cotton, C., Cronin, A., Davies, R., Davis, P., Dear, P., Dearden, F., Doggett, J., Felwell, T., Goble, A., Goodhead, I., Gwilliam, R., Hamlin, N., Harte, G., Harper, D., Hauser, R., Hornby, T., Holroyd, S., Horrocks, P., Humphrey, S., Jagels, K., James, K.D., Johnson, D., Kierkegaard, A., Knights, A., Konfortov, B., Kyes, K.S., Larke, N., Lawson, D., Lemard, N., Line, A., Maddison, M., McLean, J., Mooney, P., Mouton, S., Murphy, L., Oliver, K., Ormond, D., Price, C., Quail, M.A., Rabbinowitsch, E., Rajandream, M.A., Rutter, S., Rutherford, K.M., Sanders, M., Simmonds, M., Seeger, K., Sharp, S., Smith, R., Squares, R., Squares, S., Stevens, K., Taylor, K., Tivey, A., Unwin, L., Whitehead, S., Woodward, J., Sulston, J.E., Craig, A., Newbold, C. and Barrett, B.G.

TITLE

JOURNAL

COMMENT

FEATURES

source

Barrell, B.

Direct Submission

Submitted (24-SEP-1998) P.falciparum Genome Sequencing Consortium, The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK

On Dec 16, 1999 this sequence version replaced gi:5763807.

For more information about this sequence or the Malaria Project, see http://www.sanger.ac.uk/Projects/P_falciparum.

1. 67970

/organism="Plasmodium falciparum 3D7"

/mol_type="genomic DNA"

/isolate="3D7"

/db_xref="taxon:36329"

/chromosome="1"

/complement(join(201..257,455..586,748..804))

/gene="PFA0575c"

/note="synonym: YAL1P3.01b"

/complement(join(201..257,455..586,748..804))

/note="similar to Schizosaccharomyces pombe hypothetical protein c119.18 spbc119.18 SMALL:Q96VGL (EMBL:AL022117) (69 aa) fasta scores: E(): 1.3e-06, 39.68% id in 63 aa, and to Homo sapiens protein 15el.1 15el.1 SMALL:15el1_HUMAN (SWALL:043715) (76 aa) fasta scores: E(): 0.00046, 38% id in 50 aa, and to Mus musculus protein 15el.1 15el.1 SMALL:Q9D822 (EMBL:AK007514) (76 aa) fasta scores: E(): 0.00054, 38% id in 50 aa"

/codon_start=1

/product="hypothetical protein, conserved"

/protein_id="CAD49075.1"

/db_xref="GI:23477009"

/db_xref="SPTREMBL:Q81227"

/translation="MENSRTISQNTIRAHVANDCEKEKYLKCFNNYKNNFLKGLTQACDYVEDYDICTVINDLKKGLSLNVEK"

/complement(join(1392..1461,1752..2598,2748..2848,2990..3276))

/gene="PFA0580c"

/note="synonym: MALLIP3.01"

/complement(join(1392..1461,1752..2598,2748..2848,2990..3276))

/gene="PFA0580c"

/note="Originally annotated as Plasmodium falciparum; conserved hypothetical protein but pfam match and other evidences suggest that it is a putative deoxyribonuclease; earlier start site possible to make a 412 aa protein

pfam match to entry PF01026 TaqD Nuclease, TaqD related Nuclease, score 117.30, E-value 2.9e-31

similar to Plasmodium falciparum conserved hypothetical protein, upf0006 family maldp3_01 SMALL:Q9UN06 (EMBL:AL031746) (412 aa) fasta scores: E(): 3.7e-150, 100% id in 412 aa, and to Saccharomyces cerevisiae putative deoxyribonuclease yb1055c yb1055c or yb10512 or yb10511 SWALL:YBPS_YEAST (SWALL:P34220) (418 aa) fasta scores: E(): 2.5e-19, 30.47% id in 397 aa"

/codon_start=1

/product="TatD-like deoxyribonuclease, putative"

/protein_id="CAD49076.1"

/db_xref="GI:23477010"

/db_xref="SPTREMBL:Q81226"

/translation="KLVFHYIKYINLVFYSIFLKSNSIKIYNDLRYISTVNRKYV LQIKRSNKKHNRIMKEDNESFIDIGSLTDFGQVYNSKHEMDLQVLRAR KNNVKKIITTCALIKESLKIETVPEGLTASVGETNCEYFIDKSHSEKE IAKKEVERFETKYPKBOVENSKNNKKIKCDGKNNNNLIEKLDDTIPQKY NKDKLEYENIKKTIKYPNRTVCGELGIDPDRIPGSKYQIKYPIFQDLVGMFN LMFPLMRKCSSTPFKIIDYKFEPEKNGVYHSSTDEDIYHITVQYKNTLYING CSUKSEINAVAKKIPINDLLETDAPMGCGVKTASVYIYIDTEKRAYTLKIKIN IIKCDNTIFKRENEPNYIADIAETTVRREAVFDLFCCK"

/gene="MALLIP3.02"

/note="synonym: PFA0585w"

5005..5496

/gene="MALLIP3.02"

CDS

gene


```
/note="contains possible signal sequence. Signal peptide
predicted SignalP 2.0 HMM (Signal peptide probability
0.640, signal anchor probability 0.357) with cleavage site
probability 0.504 between residues 27 and 28. Scansegexp
hit to PS00867, Carbamoyl-phosphate synthase subdomain
signature 2
/codon_start=1
/product="hypothetical protein"
/protein_id="CAB63557.1"
/db_xref="GI:6594245"
/db_xref="GOA:G9U0N5"
/db_xref="SPTRMBL:G9U0N5"
/translation="MKLNRRFVVICPPIILFFPLNSVYLGNNRRNINFEENAK
AKRKLSGINSIKLNDGELKIKLNDKHKSTWDKSYSPISLSEKYSQTDLP
KQOEINENTKIIDRQEFYILNDEINENIATFVLENNFELYIQSPKSLIDIIQS
LNN"
/misc_feature
8179..10303
/note="Putative centromere"
14884..20352
/gene="MAL1P3.03"
/note="synonym: PFA0590w"
14884..20352
/gene="MAL1P3.03"
/note="Putative ABC transporter
pfam match to entry PF00005 ABC tran, ABC transporter,
score 45.00, E-value 1.7e-09, HMMsart hit to SM00382,
ATPases associated with a variety of cellular activities;
ScanRegexp hit to PS00211, ABC transporters family
signature
Similar to Schizosaccharomyces pombe ATP-binding cassette
transporter abc1 abc1 or spac99.12C SMALL:ABC1 SCPO
(SMALL:093337) (1427 aa) fasta scores: E(): 4.2e-08,
21.468 id in 1146 aa"
/codon_start=1
/product="ABC transporter, putative"
/protein_id="CAB63558.1"
/db_xref="GI:6594246"
/db_xref="GOA:G9U0N4"
/db_xref="SPTRMBL:G9U0N4"
/translation="MTTKYKNGYISNKKKKKSCONTSPINFLSPDMIRPLINDLIK
GDIQLPNTGRNDVPIYASKLEBNRDRIDVEDSEFSGKSNHIVYLRHCRSDASE
KRYVLYNHNILMSILKTRFRILITISYILETLIVTGDFIDYVRIELSGQCLPV
YISFLKDFVSGLVVVMIMFHLFEALHAFYFHLFTINLVSIMYLYIKINCSN
NHLONDAPFNTYRKFSQTEIDELISDELISGKNASSSGIKNNNNINDKKEVEN
DYIINIKSKTKMEKDSLNENSLPNNYINIMFSDVPEVTFVSCINLEFVPAKF
MSRYVHIKIGNSVGIAIWMISALYSAMILPEPLPSLPKSKYLIRYRRIIDMHHVL
KPKRLIOMPMSEPAKATNIPRMKMKYCKIRIYISNIGVPSISISDIDVVEFFI
YKDRLINKBEIKFTSITMPLVYKLLISNNAFPLVNNWEGIVNKRLLNINDE
LYNDIKNFYRTRYNEDYINIVDITPLQNNENITSHDGIENHRLHLYAKNKLIN
MPCYFFPHKMNHNKINIKQILSGLKNDNTWKICFOEHSKNSYNNVNSIHE
KKEEYENIHNSNSTWSEFKEKKNNETIKLENSFGLSDKNCDDHILKNINEN
IKNSMLAIIIGVSGSKAFPSIILGDPMTGNLYIENFCKMELIVYPQSWGLCFA
NIRSMILFGENBVLVYKYLIOSELNDLSTEBGDMXYINDNRLSGQGVRLCIA
BALYEHYHNHKLCTDYREKLLIQNNELIKDLINNNKNSYNNKSKLVANNPENEN
YLOKCLMDHNNKFLYILDLITSDISISKIKFSNLCEDNINISKPNDSFISNKS
TLONFLIEDLIDVQYEVNIFELIQDKTLKRGNISYERKNLNTTKSHWYSILNT
IDYRIKLDPEVLAHVKSNKNIYKFAVKGNTESVSEIDSINKETIKKKKRY
KKEHNNKNNKNNNNNNNSKNDHININNDNRNNDINIDPNSDTDSPTVSGNE
YLDIYTSNNSDKEEIVKPLKDTHEEFKSSMPVSKSSNNKNSPFKEDSSS
FKGSISETIYLFYQOVGFYLLTSVIFMLISFIDEIFVLTWMSIISKNKSHSD
TILOKOVLYLEVFLIPITSLATSGICFEMIIYNTGTSAILKVNHLISLAPLYIF
YNNNLKNIINRIRIISADPGELKRYANAPFPFPCISLILITIMRIDCFIEPVP
IILYFVFRSRCKEAKORYLSCHTPLCNISYALSNGNININIKYKTHLDVE
HYNNRFRISYFFMLINIVASLYIKIFILLTYIIMHPLASGILIKYKKNRYI
LSTLGCISFSARLGVLIKFLCDYHIKEMCCVGRLEBPKISNKAASWKNEL
NVITTOYVEKKNENISDKLSAIVEYKNOVSGISVGSODSEKRYGILEFVNVYSQ
KKIPLVNGYKYIDEPSKLNIMYALKNOIGIVSGSAGSTILISLIGENISOG
KITVEGRDIRYVHKQEDSIIIGLAQSFVYNNIRITIDYNNETDEIVHALKN
GILKGRDLKYVHKQDKSNYKIIOTKVINOSNDNTIILINDCIRIISVRLYN
RHXYKILIDEIFIRNLNNSYHDEINSPIGRKSNPYIRNHFPNNYIILSHHANT
LSCDDIYVLRKGEITYRCSYEDVKIQSELSHLEND"
/rRNA
23896..26044
/product="18s ribosomal RNA A-type"
```

```
misc_RNA
/note="18s rRNA expressed in asexual stage parasites"
26045..26438
/product="ITS1 A-type"
/note="ribosomal operon internal transcribed spacer 1
expressed in asexual stage parasites"
26439..26559
/gene="mal1.5.8rRNA"
26439..26559
/gene="mal1.5.8rRNA"
/note="5.8S ribosomal RNA"
/note="5.8S ribosomal expressed in asexual stage
parasites
hit to vivax A-type 5.8S ribosomal percent id: 88.00 none"
26560..27435
/product="ITS2 A-type"
/note="ribosomal operon internal transcribed spacer 2
expressed in asexual stage parasites"
27436..31539
/product="28S ribosomal RNA (A-type)"
/note="28S rRNA expressed during asexual development"
complement(join(31566..32476,32675..32775))
/gene="MAL1P3.04"
rRNA
gene
Query Match 3.5%; Score 168.4; DB 3; Length 67970;
Best Local Similarity 43.8%; Pred. No. 7.2e-13;
Matches 1227; Conservative 0; Mismatches 1541; Indels 34; Gaps 10;
misc_RNA
328 ATTTCGAATTTATTTACAAATCATGAGTCTTTTCGACATACATATTAACCAT 387
ATTATTTCTATTTACTAAGCGATATAGAAATGATATTTCAAAATTAATCAATAAAAAT 7716
388 CAATAACAATAATAGTACTATCTAATTAATAAAAAATCCCAATAGTCTTTAGCAAGA 447
TATATGATATATATTTCTTATTGAATGCT-----CTCATMATAAGTACTTATGATGT 7770
448 ACTACAAATATATAGACACTTAATCTTATTAATATTTCAATTTATATCTACTCA 507
7771 TCTAAATATTTATTTGAATTAAGCAATTAATATTTCTAAGAAAGGACTTTATTAATA 7830
508 TCATATTAGCAATATATCATCTGACATGATTAACAATAGTATTAACATTTCTTAAC 567
7831 ATATTCCTTATTTATATTAACATTTATTAATAATATTAAGAAATTAATTTCTTAA 7890
568 TAGATAAT-ATACATAAATATATATATATTTTGACTGTTCTTGATTAATAGGT 626
7891 AAACGATGATATTAATAAATTAATATATATATTAATTTATATATATATATATAT 7950
627 AAAAAATGAATAAGATTTTATAGGACAGATCATTTGGACAGCAATAGCAATGGAC 686
7951 ATTTCATTTATTTTATAGTTTATATATATATATATATATTTCTTATTAATTAAT 8010
687 TGTGACACATGATATATCATGTTAATCTTATGGGACGTGATCATAGCAATTTACATA 746
8011 TATAGTACATTTATATTAATAATATA-----ATTAAGATATTTTATCTTT 8060
747 TAACACTTACCGGAAAAAATTTAAATATTTACCCACATATCTCCAGACACATAT 806
TGATTTATATATATATATATTAATTAATGAATAATTAATTAATTAATTAATTAAT 8120
807 ATGCAACGACGATATAGACGAGATATATCTGACAACAATACATGATATCTTAACTA 866
8121 ATATTAATAATATATATATTAACAAATAATATATATTAATTAATTAATTAATTAAT 8180
867 TTCTCAACACATATTCGACCATATGATACAAATACAAACCAACAGCATATGCTA 926
8181 AATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8240
927 AAAATACCTCAATATTAATTTGATAGTAAGGATCGAAATTAATCATCTCAATTTCCA 986
8241 AAAATATTAATAATATATATTAATTAATTAATTAATTAATTAATTAATTAATTA 8300
987 GTGAATATGACGAAGAAATTTTCAATATATATCAATTTTATACCTCTGCTCAAGTT 1046
```


DEFINITION Drosophila melanogaster complete mitochondrial genome.
ACCESSION U37541
VERSION U37541.1 GI:1166529
KEYWORDS Mitochondrion Drosophila melanogaster (fruit fly)
SOURCE Drosophila melanogaster
ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Preygotha; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 12511 to 12682)
AUTHORS Clary,D.O., Goddard,J.M., Martin,S.C., Fauron,C.M. and Wolstenholme,D.R.
JOURNAL Drosophila mitochondrial DNA: a novel gene order
MEDLINE 83090428
PUBMED 6294611
TITLE 2 (bases 5269 to 5695)
AUTHORS Clary,D.O., Wahleithner,J.A. and Wolstenholme,D.R.
REFERENCE Transfer RNA genes in Drosophila mitochondrial DNA: related 5' flanking sequences and comparisons to mammalian mitochondrial tRNA genes
JOURNAL Nucleic Acids Res. 11 (8), 2411-2425 (1983)
MEDLINE 8320794
PUBMED 6304652
REFERENCE 3 (bases 404 to 5272)
AUTHORS de Bruijn,M.H.
TITLE Drosophila melanogaster mitochondrial DNA, a novel organization and genetic code
JOURNAL Nature 304 (5923), 234-241 (1983)
MEDLINE 83245048
PUBMED 6408489
REFERENCE 4 (bases 804 to 1778)
AUTHORS Satta,Y., Ishiwa,H. and Chigusa,S.I.
TITLE Analysis of nucleotide substitutions of mitochondrial DNAs in Drosophila melanogaster and its sibling species
JOURNAL Mol. Biol. Evol. 4 (6), 638-650 (1987)
MEDLINE 2832697
PUBMED 5 (bases 5268 to 13619)
AUTHORS Garesse,R.
TITLE Drosophila melanogaster mitochondrial DNA: gene organization and evolutionary considerations
JOURNAL Genetics 118 (4), 649-663 (1988)
MEDLINE 8812147
PUBMED 6 (bases 441 to 2967)
REFERENCE Satta,Y. and Takahata,N.
TITLE Evolution of Drosophila mitochondrial DNA and the history of the melanogaster subgroup.
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 87 (24), 9558-9562 (1990)
MEDLINE 91088557
PUBMED 7 (bases 14215 to 14512)
AUTHORS Ballard,D.W., Olsen,G.J., Faith,D.P., Odgers,W.A., Rowell,D.M. and Ackinson,P.W.
TITLE Evidence from 12S ribosomal RNA sequences that onychophorans are modified arthropods
JOURNAL Science 258 (5086), 1345-1348 (1992)
MEDLINE 9308057
PUBMED 8 (bases 14917 to 19517)
REFERENCE Lewis,D.L., Farr,C.L., Farguhar,A.T. and Kaguni,L.S.
TITLE Sequence, organization, and evolution of the A+T region of Drosophila melanogaster mitochondrial DNA
JOURNAL Mol. Biol. Evol. 11 (3), 523-538 (1994)
MEDLINE 94285822
PUBMED 9 (bases 1 to 408, 13319 to 19517)
REFERENCE Lewis,D.L., Farr,C.L. and Kaguni,L.S.
TITLE Drosophila melanogaster mitochondrial DNA: completion of the nucleotide sequence and evolutionary comparisons
JOURNAL Insect Mol. Biol. 4 (4), 263-278 (1995)
MEDLINE 96423163

PUBMED 8825764
REFERENCE 10 (bases 1 to 19517)
AUTHORS Lewis,D.L., Farr,C.L. and Kaguni,L.S.
TITLE Direct Submission
JOURNAL Submitted (03-OCT-1995) Laurie S. Kaguni, Biochemistry Department, Michigan State University, East Lansing, MI 48824-1319, USA
FEATURES
source
1. 19517
/organism="Drosophila melanogaster"
/organism="Mitochondrion"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/note="derived from new and previously submitted sequences; sequence is a composite containing sequences obtained from different Drosophila melanogaster strains"
1. .65
/product="tRNA-Ile"
complement(97. 165)
/product="tRNA-Gln"
171. .239
/product="tRNA-Phe"
240. .1265
/codon_start=1
/trna1_table=5
/product="NADH dehydrogenase subunit 2"
/protein_id="AAC47812.1"
/db_xref="GI:1166530"
/translation="MENNSSKILFTIMTIGTITLITTSNSWAGNMGEINLSFPL
LSDNNIMSTESALIKYELQVASTVLPSILMKNNNNNINESFTSMITMGL
LKGAAPFHFPPNMMEGLTMMNMLMTMOKLAPMLISYIKYLLI SVLSYIT
GAIGAGNOTSLRLNAFSSINHLGMLSLIMISEIWLIFPYSFLSVLFEMVIF
KLHLNOLFSFVNSKILKPTLPFNLISGLGPPFLGPKMLVIOQLTCLNQYFULT
LTKMSTLITTFYLRICYSAPFNNYFENNMMNANNNSINYNMIMTFPSIFGLFI
SLRFNF"
1264. .1329
/product="tRNA-Trp"
complement(1322. 1383)
/product="tRNA-Cys"
complement(1403. 1468)
/product="tRNA-Tyr"
join(1470. 1472,1474. .3009)
/note="mechanism underlying reading frame shift after first codon uncertain"
first codon start=1
/codon_start=1
/trana1_table=5
/product="cytochrome c oxidase subunit I"
/protein_id="AAC47812.2"
/db_xref="GI:7412848"
/translation="MSKQMLPSTNHDIGTLYIFGMAAGVNGNSILIRAE.GHPG
ALIGDDIYIVIVTAHAFIMFPMVPMIGGFGNMLVPLMLGAPDAFRRNNMSW
LNPALSLILVNSSTGISLDRMPLFWSVIVITALLLSLPVLAGAITWLLDRNINISF
VNFITIVINASTGISLDRMPLFWSVIVITALLLSLPVLAGAITWLLDRNINISF
LPDAGGDPILYCHLFWPFGHPEYVILILPGFGWISHLISQESKKEFTFSLGNIYIM
LAILLIGFTVAHMHFTVGDVDRAYFATMIIAVPTGIKIFSMATLHGTOLYS
PALIMAGFVPLFTVGGTLGTVLIMFISVNIITFPOHFLAGMGRRSDDVATYTN
IHWFTPLGTLLNKKKSHKSHFIMFISVNIITFPOHFLAGMGRRSDDVATYTN
IVSTIGTSLIGLIFFFITLWESLVQRQVITYIQLNSSIEWQNTPPRHSSELP
LLTN"
3012. .3077
/product="tRNA-Leu"
complement(3073. 3767)
/note="TAA stop codon is completed by the addition of 3' A residues to the mRNA"
/codon_start=1
/trana1_table=5
/product="cytochrome c oxidase subunit II"
/protein_id="AAC47813.1"
/db_xref="GI:1166532"
/translation="MSTNANGLGDSAPLMLEOLIFPHDHLILIMVTVLVGYLME
LFNNYVRLFLHGLLMTITLIPALITLIPALITLIPALITLIPALITLIPALITL
QWYSYSDSDNNNISFSDYIMIPNELMTDGRLLDYNNRYLTPNSQIRITLVTAAYI
HSWTVPALGVKVDGTPRLNQTINFFINPGLFYQCCEBICGANSFMPYIESVPVY

1463 AACCTAAAGCTGTGAGAGAACGCTACGTGTAACCAACGCTGATATCATAGCAAAAAA 1522
17603 TTTAAAAACATGATTTTATATATATATATATATATATATATATATATATATATAT 17544
1523 CTATATGAGAGACTACCAATTTGCAATATCTATATATATATATATATATATATAT 1582
17543 TTTAAAAAT 17484
1583 CTAGTAAAGCATATCAAACTATTT--ATCTATCTATCATCTCTTGATATATTCAC 1640
17483 TATTAATATATCAAT 17424
1641 ACAATGATTAATTAACGCTCAAAAATATAGATACAGCTTACACCATATATATATAT 1700
17423 AAAAAAAT 17364
1701 AGTCAATATCCATGCTTACATCC--AACTAGATGTAACATATATATATATATATAT 1758
17363 ATTTAT 17304
1759 AAGTATTAATTAAT 1818
17303 TAT 17244
1819 TAAACATTTGTATCAGATTAATATATATATATATATATATATATATATATATATAT 1871
17243 TAT 17184
1872 TAAATATATATGTTGACACATTTTCTACATCATGATATATATATATATATATATAT 1931
17183 TAAAAAT 17124
1932 ATTAAT 1982
17123 AAAAAAAT 17064
1983 CTAT 2042
17063 TAT 17004
2043 TATTTCAACATTAAT 2102
17003 AGTAT 16944
2103 GCTAT 2152
16943 GAT 16884
2163 CAGCTAT 2222
16883 TAT 16831
2223 TAT 2282
16830 TAT 16771
2283 GTAT 2342
16770 ATTTTAT 16711
2343 TAT 2402
16710 TTTTAT 16651
2403 ACAT 2459
16650 TAT 16591
2460 GCAATGTTAT 2519
16590 ATTAAT 16531
2520 AAGCTTCAATCAAAATTCAGAGCTTGTATATATATATATATATATATATATATAT 2579

16530 CTAT 16471
2580 GAGTTT--TTAAATCTGCTTTCAATTTCTTACAGAGCTTTAGTATATGTAAGAAAT 2637
16470 AGTAT 16411
2638 AACTTATGACCTTAAACATCATTAATAGGCAATATATATATATATATATATATATAT 2697
16410 AT 16351
2698 TTTTCTGAT 2757
16350 TAT 16291
2758 TATATGCTTAT 2816
16290 TAT 16231
2817 TAT 2876
16230 TTTAT 16171
2877 ATTCGAT 2936
16170 AT 16111
2937 ATCTAT 2993
16110 TCAATTTTAT 16051
2994 TGAAT 3053
16050 AGGGGAT 15991
3054 ATATTTTCTTCTGAT 3113
15990 TCAATTTTAT 15931
3114 TTTCTGAT 3173
15930 AAT 15871
3174 GTACTAT 3233
15870 TTTTAT 15812
3234 TATATGCTGAT 3293
15811 AATTTAT 15752
3294 TTTTAT 3353
15751 AAT 15692
3354 TATATGCTGAT 3412
15691 TATTTAT 15632
3413 GAGTAT 3472
15631 TAT 15572
3473 AAT 3532
15571 TAT 15512
3533 TTTTCTAT 3592
15511 TTTTAT 15455
3593 TGCCAT 3652

Tue Apr 6 07:57:22 2004

us-10-081-051-52.rge

Page 79

```
Db 15454 TTTGTAAAAAATATTTTATATATATTTTCATTTTTTTTAAAAAATTTT 15395
QY 3653 CTATATCATCATATATCATGCTCTTACCATATATCCATATATACAAATCAGGCTCT 3712
Db 15394 TTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 15335
QY 3713 TTTTCATTAAGCATCTGATCGCGATTAACATCAAGCTGATTTAGTGAATATTA 3772
Db 15334 AAAAAAATTTTATTTAGCATTTAAAGAAATTAATCATATATTTTAAAAATTTT 15275
QY 3773 TCATCACTTAATATATCTTTTATTAATTTGATTTGATATCACTCATCATCAACTAT 3832
Db 15274 TAAAAAATTTTATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAA 15217
QY 3833 ACTTCAGAAATATATCAATCTTCTGACATTTGAAAAATCCTGTCATTAATTTACCTAC 3892
Db 15216 AAAGTGAATTAATTAATTTTATTTTAAAAATTTAATTTATTTATATTTGA 15157
QY 3893 AATATCATGAATAATCTGCAACACA-AGTACTCTCTTAAGTACACTCATATATCTA 3951
Db 15156 TAAATTTATTTTGTAAAAATTAATTTTAAAAAATTTTAAAAAATTTTAAAAA 15097
QY 3952 CCACACACATATATGATCAACACCAATATTAATTTCAATTTTACTTATTAAT 4011
Db 15096 TTGTAAAAATTTAAAAATTAATTTTATTAATTTATTAATTAATTAATTAATTA 15037
QY 4012 TAAACATATTAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4069
Db 15036 TTAATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTTCTTAA 14977
QY 4070 ATATATCCAAAAAGAAAAATTAATAAATTAATAAATTAATAAATTAATAAATTA 4129
Db 14976 AAATTAATTTTATTAATAATTAATTAATTAATTAATTAATAAATAAATAAATA 14917
QY 4130 CTATATGATTAATTAATTAATAATTAATTAATTAATTAATTAATTAATTAATTA 4189
Db 14916 TTAAGTTTATTTTGGCTTAATAATTTGTTATTAATTAATTAATTAATTAATTTGT 14857
QY 4190 GTCTATAGTTTTCATATCTAAATACTGACTGAACTTATAGATTTCAATAT 4244
Db 14856 GTGATTTTATTAATTAATTAATAATTAATAATTAATTAATTAATTAATTAAT 14802
```

Search completed: April 5, 2004, 18:26:33
Job time : 18045 secs

Db 776 GCCAAGACGGCTAATAAGCGGGTTTGAATATCCAGTGGCATATGCCCAATTTTCATCAAGA 717

QY 3128 AATAAGCTACCAATGATTTGGTGTCTCTATATATTCATATAGAGAGATCTGTATGAC 3187

Db 716 AAGAGGGAACCGGCATTGTGTGTCTCAAAACGGCCCTTGCGACTGTGAGAGCACAGTA 657

QY 3188 AAGATATTTATCTACTTCCACACATCAACAATATATTAACCAATATTTATAGCTGTAGC 3247

Db 656 AATGCCCCCTTTTCATGACCAAAAAGTTCTGATTCATTTAAGTCTTTGGAAATAGCCGC 597

QY 3248 ATAGATAGATACATAGATATATTAATGAGATACATACCCCTTGATTTTTTATGTTATAGC 3307

Db 596 ATA-----TTGAGTGCATTAATAAGGGGCTTGCTGTGCACTATGGCGATGAAGACA 543

QY 3308 CTGACAACTACTCTCTCTTCCAAACCTGGGGAACCGATATGATATGATGCACTGATGTA 3367

Db 542 TGTGCAACTATATCTTTTACAGTACCCGACTGCGCACTTATATAGACATATATGATAG 483

QY 3368 GTAAGCTGCTTTATTAATCATACTTCTCAATTAAGTATTAACGAGGGAGTTAACGACTAT 3427

Db 482 CGAATAAGGCGGCATATATGCGATTAACCTCTTGATGCGAGGGCTTCCGCAATCAAG 423

QY 3428 TCATAT---ATCCCTAATAGCTGATTTCACTCATCTCATTTCTCTAGTATTAACAAG 3484

Db 422 TCAAGACAGATTTGAAGGTTAGTTTCTGTGTTATAGGTTGTTTTTGTCTCTGATAGTC 363

QY 3485 TCTATAGCTCTCTTTACACTAACTTAAATCTTCTCTCTGTAAGAGGCTTTTCTATATTA 3544

Db 362 GTAATACACGATGAATGAGTGAAGTATCATCAATATCAAAAAGGCTTGGGAGATTA 303

QY 3545 TCATTAAGCAACCATATGACAGACTTTACAGAGTGGCAATATTTCCATGCCAATATAC 3604

Db 302 TCAATATCCCCCTTGTGATATAGCAATTTACCGGGCATCAAGATCAAGATGCCCTGTCAAG 243

QY 3605 ATAATAACAGGCAATATAGGATACCTTTCTTTAAGCTTTTCCAGTACACTTAATCATCA 3664

Db 242 ATAATTAACCGGCAACGTAAGGGTGTGTCTCTTTAACGCTTTTAATATGAGATTAACCATCA 183

QY 3665 ATATCAAG 3671

Db 182 ATATCAAG 176

RESULT 2

US-09-328-352-3206/c

Sequence 3206, Application US/09328352

Patent No. 6562958

GENERAL INFORMATION:

APPLICANT: Gary L. Bretton et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

FILE REFERENCE: GTC99-03PA

CURRENT APPLICATION NUMBER: US/09/328,352

CURRENT FILING DATE: 1999-06-04

NUMBER OF SEQ ID NOS: 8252

SEQ ID NO 3206

LENGTH: 1557

TYPE: DNA

ORGANISM: *Acinetobacter baumannii*

US-09-328-352-3206

	Query Match	Similarity	Score	DB	Length
Best Local	53.1%	48.3%	160	DB 4	1557
Matches	553	0	Mismatches	575	Indels 18, Gaps 3
QY	2659	TTGTGCTGAATCATCTCTCTAGAGATTTCATATTAATTAATTCATTCTATTAACATTACGT	2755		
DB	1208	TCCTTGTGATTAACCTCACGACCAATTAATCATGACGGTTAAACCAACGACAGATTTTCA	1149		
QY	2759	AATTGCTTAAGTTACCTGGCCATTTCATATGATCGATTCATTAATTAAGTTTCATCCTT	2818		
DB	1148	AGCTACGACATTTTCAGGCGCATGTGTAATTTGTTTCATATATGTCGTTGTTCATGCGT	1089		

QY	2819	AATACATGAGTACACAAACCTATTTTTCACAGATGCTATTCATTAATACCAACAT	2878
Db	1088	AAAATCTTTGACCTTACGCCCTAGCTCTTTCACCCGCGACCAAGAAATGTGGGCCAC	1029
QY	2879	TCCGATATCTGACAAATATTTCTACTMAAGCGGTACTCTMAATTGGMAAGACATTTAT	2938
Db	1028	ATCGGAATGTCTTACGTCCGATGGCGAAGCTCGGAATATGAATTCGAATPAACATTAGA	969
QY	2939	CTAATATATTAATCCTCACAACAAACCTACGAGCTTTTCTTCACTTCAATATGTTGAA	2998
Db	968	CGGTGATATAGTCTTACACGGAACGACCTTCATCCACAGTTTCTTAACTCTTATGG	909
QY	2999	GAAGACACAAATATTCCTACATCTATCTACTPAACAGAACTCTACTATTTTCCATATATT	3058
Db	908	GTTCCCGGACCAATTCTGACATCTACTTAAACGGGTATATGTCACCAACATCGGTAAAC	849
QY	3059	TTTCCCTCCGTATATATCTGAGAACTTAAATGGGTATCGATGCTAAATCTGTACT	3118
Db	848	TCACCATCTGGACACACACCTBACACATCGATTGTGTTCMAAGGCACTACCAAT	789
QY	3119	TCATCTATTAATPAACGTACACATGATTTGCTGCTCTATTAATCCATATAGAGAGTACT	3178
Db	788	TCATCAAGAAATAGTGTTCGCCATTTGCTGTTCMAAGGGCCCTGATGTGGGTGTT	729
QY	3179	CTATGACACATATATATTACTTCTCTCACTACCAATATATTAACCAATATATTA	3238
Db	728	GCACCCGTGA-----AGCACCTTTTCATGACCAATATATTCGTTTGATTAATCT	675
QY	3239	GCTGTAGACATAGATGAGTACATAGATATTAATGAGATATCATACCCCTTGATTTT	3298
Db	674	TTAGAAATGTGCGCATAGTTTAAAGCAATAAAGGCTTACGCGCTCGTGGATATGTTA	615
QY	3299	TGATATTAGCCATGAACTATCTCTCTTCCAAACCTGGCGAACAGATATGATATGCA	3358
Db	614	TGTAAGCATGTGACCAACGATTTCTTACCTGTGCTATTCACATTAATTAACACTGA	555
QY	3359	CTGATGTAGTAGTCTTATTTATATCATATCTTCAATTAAGATATGACAGGGACTTA	3418
Db	554	ATATGCGCTGAGATTAACGACCAATCGCTCGAATAACCTTGTGCATAGACAGGATTCG	495
QY	3419	CGGATATTGATATATCTCA-----ATGCTATTTCAACTCATGATTTTCTCTAGT	3472
Db	494	CCAAATATTTCTGTGATGTAGTGGTGAAGCTGTTTTGTGCGCTTCTTGCTGTGAAT	435
QY	3473	AATCTACAGACTCTATAGCTCTCTTTTACAACTAATCTTAATCTTCTCTCTGTAAGGC	3532
Db	434	TTATGTATATGTAATAATCGCTCGATTGACTMAAGCTMAAGCTCATATATGCAATGGT	375
QY	3533	TTTTCTATATATATATAGACCCATATGACATATGACACTTACAGAGTGGCAATTTTACA	3592
Db	374	TTTGGTAATATCTCAATAGCACACGCTGATAGCTTGATATACAGAGATCTTAATCCGA	315
QY	3593	TGCCCACTAATCATATATACAGGCAATATAGAGTACTTTCTTTAAGCTTTCCAGTACA	3652
Db	314	TGCGCTGTCATATGATATACAGGCAAGTCGGGGTGAAGATTTTTCACCTTATGATAGAA	255
QY	3653	CTTATCATCATATATGATCCCTCTTAACCATATATCAATATATACACATCAGGCTCT	3712
Db	254	GTTAAACCATTAH-----CTGGCATATGATATGATATGATAGACATGGGGCGCA	201
QY	3713	TTTTCATPAAGCCATCTTGATGGCGATPAACATCAACTGCTAATTTTAGTACATATTA	3772
Db	200	TCATATGTAACGTTCAAGCTGCTGTGAGACTTCTTCAAAATGGTGACCTCAATCTT	141
QY	3773	TCATACATTAATATCTTTATATAGATTTGTGATATCAACTGATCATCAACAATAT	3832
Db	140	TCTTCTTTAAAGTCTTTTCAAGTACCATGCAATGGGCGCATGCTATCATATAACCAT	81
QY	3833	ACTTCA 3838	
Db	80	ATTTTA 75	

Db 80 ATT TTA 75

RESULT 3
US-09-107-532A-350
Sequence 350, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Achilleo, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 350:
SEQUENCE CHARACTERISTICS:
LENGTH: 1071 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc_feature
LOCATION: (B) LOCATION 1...1071
SEQUENCE DESCRIPTION: SEQ ID NO: 350:
US-09-107-532A-350

Query Match 3.3%; Score 155; DB 4; Length 1071;
Best Local Similarity 49.0%; Pred. No. 3.9e-20;
Matches 504; Conservative 0; Mismatches 515; Indels 10; Gaps 3;

QY 584 AATAATATATATTTGACTGCTTCTTGATATATATTTAGTAAAAATGAAATCAGT 643
DB 10 ACTAATATGAAATATATATATTTATTTAGAGAAATATCCATGAAACAAAAATGCA 69
QY 644 ATTTAGTGCAGATCATTGGGACAGCAATACCAATGACCTGTAGACATGATATA 703
DB 70 GTTCTGTGCTGATCTGGGGACATGCTTTAGCAAAACATTGGAGAAAACGGGCAT 129
QY 704 TCAGTTAATCTATGAGGACGTGATCATGAAATATTTACATATATAACCTTACCGAAA 763
DB 130 GATGCGCATATTTGGGAAATGTACCTGACAGATGATGAAATCAATAGTATCATACA 189
QY 764 AATTTAAATATTTTACCAATATCATTTACCAACAACATATATGCAACCGAATATA 823
DB 190 AATGACATTTTAACTGATTTGACGATCCCAAGATCGATTATCGATATTAAGAAATTG 249

QY 824 GACGAGATATATCTGACACAAATAC--ATGATATATCTTAATCTTCAACACAT 881
DB 250 GCAGAGCAGTGTATACGACAGATGATATTTATTTGTTCCAAACAAAGCATTCGT 309
QY 882 TAGCAGCATATGTACACAAATATACACAAAGACATATGTAAAAATCTCCATAT 941
DB 310 TAGTGCACGAACTTATTTGTAATATGTAATCTAAGCCATATATACAGCAAGT 369
QY 942 TAATTTAGTAAAGATATCGAAATTAATCATCACTCAAAATTTCCAGTAAATAGCAAG 1001
DB 370 AAGGATAGAAACAGATATACGATTAAGAAATCCAGAAAGTTTGGCTGAAAGAAATCCA 429
QY 1002 AATTTTCAATATATATCCATTTTATATCTCTGCTGGCCAGTTTGGCTAAAGAAATTG 1061
DB 430 GAA--GAAAAAAGACAGCAATGTGTATATATCTGCTAGCCATGCTGAGAAAGTGG 487
QY 1062 CAGAACTCTTCTTGTAGTATAGTACTGCTGCTGTGATATATTAAGAACTTGTGATCAT 1121
DB 488 CAGTACATGATATACAAACATTAAGCTGCAAGATGATCAAAAGACAGCCTATG 547
QY 1122 TGATGAAACATATAGTATATGATGTTCTAAATAATATACATCAAGATTTATAGTG 1181
DB 548 TACAAAACTGTCTGATGATATATTTACAGATTTATACAAATATGATGTGATGGCG 607
QY 1182 TACAGATTGAGCTGCATTAAAGAACATATATGCAATTGCACTGTGATATATCGCTGGA 1241
DB 608 TAGAACTGTGTGACACTGAAAAATATCATTTGCAATTTGTCGGCCGCTATTCATGAT 667
QY 1242 AAAATTTAGTATATATCTGCTGTCTACTGTTATATCAATGAAGCATGATGAATTA 1301
DB 668 TAGATTCGAGATGATCAAAAGACAGCAATCATGACCTCGAGATGCGCAAAATCAACC 727
QY 1302 CACTATATATGCAAAAAATCATTCATATAGATCTTCATACATTAATGTCATCATGTC 1361
DB 728 GTCTGTGTGCGCATGGGT-----GCCAATCTCTGACATTTATGCTGTGAGTGTG 781
QY 1362 TTGAGATCTATATATTAATGATGACACAGAAATTCACGCAATATGCTTTGGACTAG 1421
DB 782 TGGGGACTGTATCTGTACTGTTCAAGTGTTCATCTCGCAATTTGGCGTCCGTAATC 841
QY 1422 AATATGAAAAAGTAAAT 1481
DB 842 TGTTAGGACAGGACATATGCTGGAAGATTTAGAAAAACATGGGAATGCTGTGAGG 901
QY 1482 GAACAGTACTGTAAACCACTGATATCATTTAGCAAAAAAATTAATATAGACTACCA 1541
DB 902 GTGTGGCAACGACAAAGACAGATGAGCTAGCTCAACAGCTAGTGTGAGATGCCAA 961
QY 1542 TTTCATATCTATTTTCAATTTATTTACATGAGAAATATATCTACTATATTAACCATAT 1601
DB 962 TTACTCAACATCTATATATGTTTGTATATGAGAAAGATATCAAAAAACAGCAAAAG 1021
QY 1602 ACATATAT 1610
DB 1022 AATCATGT 1030

RESULT 4
US-08-956-171E-513/C
Sequence 513, Application US/08956171E
Patent No. 6593114
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
Gail H. Choi
Patrick S. Dillon
Craig A. Rosen
Steven C. Barash
Michael R. Fannon
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5256
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue

CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,171E
FILING DATE: 20-Oct-1997
CLASSIFICATION: <Unknown>
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark U. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 513:
SEQUENCE CHARACTERISTICS:
LENGTH: 918 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 513:
US-08-956-171E-513
Query Match 3.2%; Score 154.2; DB 4; Length 918;
Best Local Similarity 53.6%; Pred. No. 5.4e-20;
Matches 346; Conservative 0; Mismatches 293; Indels 6; Gaps 1;
DB 1022 ATTATTAATCTCTGCTGCTCAAGTTTGGCTAAAGAAATGCGAAGATCTTCTGTGTGT 1081
DB 732 ATTGGCGCTGTGTCAGGCGGCTCAATGCGAAGAGTTGTATGTCAGCAACCACTACA 673
DB 1082 ATAGTACTTGCTGTGATATAAAGAACTTGTGTAATCATTTGATAGAAACAATAAGTAT 1141
DB 672 GTTGCGCTTCATCAAAAGATTAAGTATGTAATTAACGCAAGATTTATTTATGTAAT 613
DB 1142 GATGCTCTAAATATATATCAATCAAGATATATAGTGTACAGATTGGAGCTGCATTA 1201
DB 612 GATTATTTGGGTGTACACGAATGATGACTGTGATGGTGTGAACTTGTGTGCTGATG 553
DB 1202 AAGACATAATTTGCAATTGCGATGGAATTAATGCTGGAAGAAATTTAGTATATATGCT 1261
DB 552 AAAAATATCATCCAGAGCAAGTGAATGATGCTGGAATGGCTACGCTAGATATATGCA 493
DB 1262 GTTGCTACTGTATTAATCAAGCAAGTGAATTAATTAACATAATATATACCAAAAAAT 1321
DB 492 AAAGCTGCAATTAATGACTGCTGCTTACCGGAATTAATGAGA-----TTAGGTGAAG 439
DB 1322 CATTCATATAGATCTTCAATCAATTAATTTGCTCATCATGCTTGGAGATCTAATATTAACA 1381
DB 438 TTAGGTGCGCATCTTATGACATTTTATAGTTTATAGTGTATGATGATGATGATGATGAT 379
DB 1382 TGTACACAGAAATTTACCGCAATATGCTTTGGCTAGCTAGAAATAGAAAGAGTAGAAT 1441
DB 378 TGACATCAACACATTTCTGGAATTTCACTATGATATTAATTAATTTGAGAAAGTAGTCA 319
DB 1442 ATAAATCAATTAATAGTCAACACTAAAGCTTGTGGAAGAAACAGTACTGTATTAACCA 1501
DB 318 ATGATCAAGCATTTATCTGAAATGAAATATGATGTTGTTGAAGATATTAATCAACTAAATCA 259
DB 1502 CTGATATATCTTACCAAAAAAATTAATGTAAGAACTCAATTTGCAATATCTATTTACAAAT 1561
DB 258 GTTATCATTTAGCTAAAGAAAAAATGAGATATGCAATTTACAAATGCAATTAATATAGA 199

1562 TTATACATGAGATATATATCACTAGATTAAGCCATATCAACATATATATCTAGTCAATC 1621
DB 198 GTATTAATTTAAATATCTCATGTAAGTAAGAAATGCTTAAAGATTTATGAGCGCCGATTA 139
DB 1622 ATCTTGTGATTAATTTCAACCAATGATTAATTAATTAAGCGTCAAAAAT 1666
DB 138 AAATCTGAATTAATTAATGATTAATTTACATATATACGCTAATTAAT 94
RESULT 5
US-08-107-532A-1399
Sequence 1399, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOMER THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Denek
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781) 893-5007
TELEFAX: (781) 893-8277
INFORMATION FOR SEQ ID NO: 1399:
SEQUENCE CHARACTERISTICS:
LENGTH: 1062 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...1062
SEQUENCE DESCRIPTION: SEQ ID NO: 1399:
US-09-107-532A-1399
Query Match 2.9%; Score 139.8; DB 4; Length 1062;
Best Local Similarity 49.6%; Pred. No. 2.5e-17;
Matches 484; Conservative 0; Mismatches 467; Indels 24; Gaps 4;
DB 645 TTTAGGTGAGATCATTTGGCAACAGCAATGCAATGCTGTCAGCATGATATAT 704
DB 56 TTTAGGACCAAGGCTCTTGCGGTACTCTTGGCAAGTCTGCGGAAACGCAATG 115
DB 705 CAGTTAATTAATGCGGCGGATCATGAATATTTACACATATTAACACTTACGAAAA 764

Tue Apr 6 07:57:22 2004

us-10-081-051-52.rn1

Page 5

116 AGCTTCGATTTGGGACATCGTCTTCAAGTAGAAGAAATCAACCCACAGCA 175
176 ATTAAATATTACCCACATATCATCTACAGACAAATATAGCAACCGCAATTAG 824
176 ATCAAGCTATTACAGATCATTTATACCGGATTCATCCAGATTCACCAATAGG 235
825 ACGAATATTATCTGACAAATATAGATATTTATCTTACTATCTTCAACAGATTA 884
236 AAGATGAGTAAAGAGCGAGATGCG--GTCCTATTCGCTATTCACGAAAGCAATAC 292
885 GCAACCATATGACAAATATCAACAAAGAGATATGCTGTAATATCTCAATTTAA 944
293 GTTCTGTTGCAAAAGACTAGTCTTAA-----TTGATATCAAGCCAGTATCATCC 346
945 TTGTAGTAAAGTATCGAATTTACATCATCAAAATTTCCAGTAAATAGCAGAGA-- 1002
347 ATGCAAGTAAAGATTAAGAAAGGAGAAACATTAAGATCTTACATCTGTTAGAG 406
1003 -----AATTTACATATATCAATTTTATCTCTGCTGCTCAAGTTTGTAAAG 1055
407 AATCCAGCAGAAAAAAGAGAGTGTGTTGTTCTTCTGACCAAGCCATGAGAG 466
1056 AATTCAGACATCTCTCTGATATAGTATCTGCTGCTGATATTAAGAACTGAGT 1115
467 AAGTACCAAAATGACATCACTACAGATCAGCTGCAAGTACAGATGCAAAAGAGCA 526
1116 AATCATTTAGTAAACATATAGTATGATGTTCTTAAATATATATACATCAAGATTA 1175
527 AATATGATACAAAGCTGTTCAATGATATATTTTCGATCTATACCATCTGATGGA 586
1176 TAGGTGATAGATTTGAGAGTGCATTAAGAAACATATTTGCAATTCAGTGAATATCG 1235
587 TCGGTGATAGATTTGAGAGTGCATTTGAAATATTTATTTGCTATGCTGCTATTC 646
1236 CTGGAATAATTTAGTAAATATGCTGTTGCTGCTGTTATTAATCAAGGATGAATGAA 1295
647 ACGGCTTAGGTTTGGCGATTAATGCAAAAGGAGCAATCATGCGCGGATTTAGCTGGA 706
1296 TTTAAACATATATATAGTAAATATGATGATGATCTTCAATCTTATTTGCTCAT 1355
707 TACGCGTTAGTGTAGATGAGTGGTCC-----ACCGTGTGCTTTATTCGATTA 760
1356 CATGCTTGGAGATCTATATTTATGATGATGATGATGATGATGATGATGATGAT 1415
761 GGGAAATAGGCGATCTAGTCTTACGCAAGTGTCTTCTGGAAGTGGGCGGAG 820
1416 GACTGAAATAGGAAAGGATGAAATATTAATATATATGATGATGATGATGATGAT 1475
821 GGAATCTTGTAGGAAAGGAGCGCCCTTGAAGAGTGTCTGCAATTTGGGAATGGCG 880
1476 TTGAAGGAAACCGATCTGTAACCACTGATATCTTACGAAATTTATATGAGAAC 1535
881 TTGAAGGATGATATCTACCGAGAGGATGATGATGATGATGATGATGATGATGAT 940
1536 TACCAATTTGATATCTATTTTCAATTTATGATGATGATGATGATGATGATGATGAT 1595
941 TCCCATCATCTCAAAATCTATGAAAGTGTGATGAAACAAATTTCAAAACAGCGG 1000
1596 TATCAACATATTTAT 1610
1001 CAAAGAAATCATGT 1015

RESULT 6
US-09-557-884-1
; Sequence 1, Application US/09557884
; Patent No. 6506581
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: The Nucleotide sequence of
; the Haemophilus influenzae Rd Genome, Fragments
; thereof, and Uses Thereof
NUMBER OF SEQUENCES: 1

CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2 inch diskette
COMPUTER: Dell Pentium
OPERATING SYSTEM: MS DOS V6.22
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/557,884
FILING DATE: 25-Apr-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/476,102
FILING DATE: JUN-5-1995
ATTORNEY/AGENT INFORMATION:
NAME: Michelle S. Marks
REGISTRATION NUMBER: 41,971
REFERENCE/DOCKET NUMBER: PB186P3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-309-8504
TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1830121 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-557-884-1
Query Match 2.9%; Score 136.8; DB 4; Length 1830121;
Best Local Similarity 50.4%; Pred. No. 3.1e-16;
Matches 476; Conservative 0; Mismatches 447; Indels 21; Gaps 5;
620 AATGATTAATAAATGAAATACATATTTAGTGTGAGATCATTTGGCAGCAATAGCA 679
635171 ATAACTTGGCAACACCAATACATGCTAGGTCGGGATCTTATGAACTGCGTGGCA 635230
680 ATGCACTGTCAGACATGATATATGATTAATTTAGTGTGAGATCATTTGGCAGCAATAGCA 739
635231 ATTACCTTTCTGCGATGTTCTCCACTACTTATGGGGGACAAATCCCGCCATATC 635290
740 ACAATATTAACCTTACCGAAAAATTTAAATATTTAACCACATATCATCTACAGAC 799
635291 GCAAAATGAGACAAAGCAAAATTTATGTTTTTCCCGATGTCATCTTCCCTGAG 635350
800 AACATATGCAACCGCAATATAGACAGATATTTGACACAAATCATGATATATTC 859
635351 GATCTTCATTTAGAAATGATATCTTGCAAGCAAT---GGAATATTTCAAGATATTTTA 635407
860 TTAATATTTCTTACACAAATTTAGCAGCATTTGTATGACAAATATACACAAACGAT 919
635408 ATCGTGTGCAAGCATCTTCTGTTGTAATTTCTATATAAT-----CAAACCGCAC 635461
920 ATGTGTAATAATCTGCAATATTTATTTATGTAAGATATGAA---ATTGATCATGCG 976
635462 TTTAAAGCCCATACCGATTTATTTTGGGCAAAAGTTTGGAGCTTAACACAGGCGCT 635521
977 AATTTCCAGTAAATAGCAGAAAGAAATTTTACATATATATTCATTTTATCTCTCT 1036
635522 TTAATCAAAACGTAATGTAAGAAACATCGGAGACATATATCTAGCTGATCTTCT 635581
1037 GTTCCAGTTTGTCTTAAAGAAATTTGAGAAATCTTCTGTTAGTATGATGATGATGAT 1096
635582 GGGCCCACTTTTGGCAAAAGAAATTTGCGAAGTCTTCCATCGCTTTTACGCTGCTGCC 635641
1097 GATATATTAAGAACTTG---GTGAATCATTTGATGAAGAAATATGATGATGATGAT 1153

Db 635642 AATAACGAGCAATTCGACGGAAATTCAGTCTGATTCATTGATGTAAGGTTCCGA 635701
 QY 1154 AATAATACCATCAAGATATTTATAGTGTACAGATTGGAGCTGCATTAAGAACAATATT 1213
 Db 635702 GTTTATATAATTCGATATGATGCGCTTCAACTTGAGGAGCAATTAATAATGTGATC 635761
 QY 1214 GCATTGCAATGTAATTCGCTGAAAAAATTTAGTAATATGCTGTGCTGCT 1273
 Db 635762 GCATTGTGCGGATTTTACAGCGTATGAGATTGGGCAATCTCCACAGACACG 635821
 QY 1274 AATACTAAGGCAATGAATGAATTAACAATATATATAGCAAAAAATCATTCAATAGAT 1333
 Db 635822 ATTACTGAGGTATTTGAGAAATTAACCCGCTAGTATCTC-----GCTCGGTGCAAT 635875
 QY 1334 CTTCATCATTAATTTGCTCATCATGCTTGGAGATCTAATATTAATCATGTCAACAGAA 1393
 Db 635876 ACCAATACATTTATGGGAATGCTGATTTGGCGATTTAGTCTTACCCGATAT 635935
 QY 1394 CATTACGCAATATGCTTTTGGACTAGAAATAGAAAAGTGAATAATTAATACATTA 1453
 Db 635936 CAATCGGCAATGAGCTTTTGGATTAATGCTAGTAAGATTAATGATGCCAAATAGCT 635995
 QY 1454 ATAGATCAACAACCTTAAGCTTTTGAAGGACAGTACTGTAAACCCTGATATCATTA 1513
 Db 635996 ATGGAATAATTCGGGCAAGTCGTGAAGGTTTATTAATACAAAGAACCTTATTTGCTT 636055
 QY 1514 GCAAAAAAATTAATGTAAGTACCAATTTGCATATCTATTTA 1557
 Db 636056 GCAAAAGACAGGTGTGAATGCGATTAACGAACAAATTTA 636099

RESULT 7

US-09-643-990A-1
 Sequence 1, Application US/09643990A
 Patent No. 6528289

GENERAL INFORMATION:

APPLICANT: Robert D. Fleischmann
 Mark D. Adams
 Owen White
 Hamilton O. Smith

TITLE OF INVENTION: The Nucleotide sequence of
 the Haemophilus influenzae Rd Genome, Fragments
 Thereof, and Uses Thereof

NUMBER OF SEQUENCES: 1
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Human Genome Sciences, Inc.
 STREET: 9410 Key West Avenue
 CITY: Rockville,
 STATE: MD
 COUNTRY: USA
 ZIP: 20850

COMPUTER READABLE FORM:
 MEDIUM TYPE: 3 1/2 inch diskette
 COMPUTER: Dell Pentium
 OPERATING SYSTEM: MS DOS v6.22

SOFTWARE: ASCII Text
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/643,990A
 FILING DATE: 23-Aug-2000

CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/487,429
 FILING DATE: 1995-06-07

APPLICATION NUMBER: 08/426,787
 FILING DATE: 1995-04-21

ATTORNEY/AGENT INFORMATION:
 NAME: Kenley K. Hoover
 REGISTRATION NUMBER: 40,302

REFERENCE/DOCKET NUMBER: PB186P1C1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 301-610-5790
 TELEFAX: 310-309-8439

INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1830121 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 US-09-643-990A-1

Query Match 2.9%; Score 136.8; DB 4; Length 1830121;
 Best Local Similarity 50.4%; Pred. No. 3.1e-16;
 Matches 476; Conservative 0; Mismatches 447; Indels 21; Gaps 5;

QY 620 AATAGTAAATAAGAAATCAGATTTTGTAGTGCAGATCATTTGGCAGACAAATAGCA 679
 Db 635171 AATACCTCGCAACACCAATCATCTGTGCTAGTGGGAGCTTATGGAATGGGCTGGCA 635230
 QY 680 ATTGCATCTGACGACATGCTATATCATTTAATCTTATGGGAGCTGATCATAGAAATTT 739
 Db 635231 ATTACTTTTCTCGCAATGCTTCCCACTCATTTATGGGGGCAATCCGCCCATATC 635290
 QY 740 ACACATATTAACACTTCCGAAAAAATTTAAATATTATCCACATATCATACAGAC 799
 Db 635291 GCACAAATGCAAGACAGAACGACAAATATATGCTTTTGGCGGATGTCATCTTCCGAG 635350
 QY 800 AACATATATGCAACCAATATATAGCAAGATATATCTGACAAACAATACATGATTTATC 859
 Db 635351 GATCTTCAATTTAAGAAATATCTTGACAGACAAAT---GGAATATCTCAAGATATTTTA 635407
 QY 860 TTACTATTTCTCAGACAAATATAGCAAGATATGACAAATATGACAAACAACAGCAT 919
 Db 635408 ATCGTGGCGCAACCCATGCTTTCGGTGAATTTCTTAATAAT-----CAAAACGAC 635461
 QY 920 ATGTGTAAATAATCTCCATATTAATTTAGTAAGATATGAA---ATTACATCATC 976
 Db 635462 TTAAGGCCATCATCCGATTAATTTGGCAACAAAGTTTGGACGTACACAGGGCGT 635521
 QY 977 AATTTTCCAGTAAATAGCAAGAAATTTTACAAATATATCAATTTTATATCTCT 1036
 Db 635522 TTACTACAAACCGAGTGAAGAACTCGGAACGCAATATCCTAGCTGATCTTCT 635581
 QY 1037 GGTCCAGTTTGTGTAAGAAATTCAGAAATCTTCTTGTATAGTACTTGTCTGT 1096
 Db 635582 GGCCCACTTTTGCAAAAGAAATTTAGCCAAAGTCTACATCAGATATTAAGCTGTGCGC 635641
 QY 1097 GATTAATAAGACTTG---GTGAATCTTGTATAGAAACAATATGATATGATTTCTTAAA 1153
 Db 635642 AATAACGCAATTCGACGGAAATTTCACTGCTGATTCATTTGATTAAGGTTTCCGA 635701
 QY 1154 AATAATACCATCAAGATATTTATAGTGTACAGATTGGAGCTGCATTAAGAACAATATT 1213
 Db 635702 GTTTATATAATTCGATATGATGCGCTTCAACTTGAGGAGCAATTAATAATGTGATC 635761
 QY 1214 GCATTGCAATGTAATTCGCTGAAAAAATTTAGTAATATGCTGTGCTGCT 1273
 Db 635762 GCATTGTGCGGATTTTACAGCGTATGAGATTGGGCAATCTCCACAGACACG 635821
 QY 1274 AATACTAAGGCAATGAATGAATTAACAATATATATAGCAAAAAATCATTCAATAGAT 1333
 Db 635822 ATTACTGAGGTATTTGAGAAATTAACCCGCTAGTATCTC-----GCTCGGTGCAAT 635875
 QY 1334 CTTCATCATTAATTTGCTCATCATGCTTGGAGATCTAATATTAATCATGTCAACAGAA 1393
 Db 635876 ACCAATACATTTATGGGAATGCTGATTTGGCGATTTAGTCTTACCCGATAT 635935
 QY 1394 CATTACGCAATATGCTTTTGGACTAGAAATAGAAAAGTGAATAATTAATACATTA 1453
 Db 635936 CAATCGGCAATGAGCTTTTGGATTAATGCTAGTAAGATTAATGATGCCAAATAGCT 635995
 QY 1454 ATAGATCAACAACCTTAAGCTTTTGAAGGACAGTACTGTAAACCCTGATATCATTA 1513
 Db 635996 ATGGAATAATTCGGGCAAGTCGTGAAGGTTTATTAATACAAAGAACCTTATTTGCTT 636055

OY 1514 GCAAAAACTTATGTAGAACTACCAATTGTCATCTATT 1557
DB 636056 GCACAAAGACAGGTGTGAATGCCATTACAGAAATTTA 636099

RESULT 8
US-09-134-000C-1318
; Sequence 1318, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1318
; LENGTH: 999
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-134-000C-1318

Query Match 2.8%; Score 134.6; DB 4; Length 999;
Best Local Similarity 48.5%; Pred. No. 2.3e-16;
Matches 465; Conservative 0; Mismatches 484; Indels 10; Gaps 3;

OY 654 CAGATCATTTGGGACGACGATGACATTTGACGTCGTCGACGATGATTCAGTTACT 713
DB 5 CAGGCTCTGGGGACGCTTTAGTTCAGTATTAGCCGAAATGACGACGAGTCTGTA 64
OY 714 TATGGGACGTCATCAGAAATTTACATATAACCTTACCGAAAAATTTAAAT 773
DB 65 TTTGGGAAATTAACGACGAAATGATGAATTAATAACACACGACGAAACATT 124
OY 774 ATTACCCATATCATCTACGACGACATATATGACACGACGATATGACGAAAT 833
DB 125 ATTACGAGATTTATCTCTCACTTCAATTCAGGACGACGATTTAGCCACTGCT 184
OY 834 TATCGACACATATCATCTATCTTA--ACTATCTACACAAATTTAGCCACAT 891
DB 185 TAGTAGTGTGATGCGGTACTTTGTGTGACCACTAAAGCATTCGTCAGTAC 244
OY 892 ATGTACCAATATACGACGACGATGTGTGTAATAATCTCCATATTAATTTGTA 951
DB 245 AAGAGTGGCCCACTTAAACAAACCAATATTATTCATGACGAAAGGCTCG 304
OY 952 TAAAGTATGAAATTCATCACTCAATTTCCAGTGAATGCGAAGAAATTTTACA 1011
DB 305 AACAGGACATTAAGCGATTTACAGAGTATAGCGAAG--ATTCCACAGAAA 362
OY 1012 ATATATTCATTTTAACTCTCTGTCGCAAGTTTGTAAAGAAATTTGACAGCAT 1071
DB 363 ACGGCAAGGATTTGTTTATCTGTCAGGCAAGCAAGAAAGTCCCTGTCATGA 422
OY 1072 TCCCTAGTATATCTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1131
DB 423 CATTAACCATTTACGACGACGACGAAATTTAGCAGCGCTGTATGTGCAAGATT 482
OY 1132 AATAGTATGATTTTAAATTAATATATATATATATATATATATATATATATAT 1191
DB 483 GTTATATATGATTTTAAATTTATATATATATATATATATATATATATATATAT 542
OY 1192 AGCTGCAATTAAGAACTAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCA 1251
DB 543 GGGGCTTTAAATTTTATTTGTTGGGCGCGGCGGCGGCGGCGGCGGCGGCGG 602
OY 1252 TAAATATGCTGTCTACTGTTATATATATATATATATATATATATATATATAT 1311
DB 603 TGAATATGCAAGGCAATTTATATATATATATATATATATATATATATATATAT 662

OY 1312 AGCAAAATCATTCAT 1371
DB 663 TCGGATGGG-----AGCAAAATCATTCATATATATATATATATATATATATATAT 716
OY 1372 AATATTAATGATGACGACGACGACGACGACGACGACGACGACGACGACGACG 1431
DB 717 AATGTTATCTGTAAGTGTGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTG 776
OY 1432 AGTGAAT 1491
DB 777 AGGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 836
OY 1492 TGTAAACCATGAT 1551
DB 837 AACTAAAGCAGCTATGATGATGATGATGATGATGATGATGATGATGATGATG 896
OY 1552 TATTTACATTTAT 1610
DB 897 AATCTACATGCTCTGTAATATGACGATGATGATGATGATGATGATGATGATG 955

RESULT 9
US-09-134-001C-300
; Sequence 300, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 300
; LENGTH: 1059
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-300

Query Match 2.5%; Score 121; DB 4; Length 1059;
Best Local Similarity 47.6%; Pred. No. 7.4e-14;
Matches 460; Conservative 0; Mismatches 495; Indels 12; Gaps 3;

OY 628 AAAAATGAAATCAATTTTATGATGATGATGATGATGATGATGATGATGATGAT 687
DB 60 AATGAGAAATTAAGTTTGTATGATGATGATGATGATGATGATGATGATGATGAT 119
OY 688 GTGACGACATGAT 747
DB 120 AGCTCAAAATGCTATGATGATGATGATGATGATGATGATGATGATGATGATG 179
OY 748 AAACACTTACGAAAAATTTAAATATATATATATATATATATATATATATATATAT 807
DB 180 TAAACACATGATATGAAACAAATATATATATATATATATATATATATATATATAT 239
OY 808 TGCAACGACATATGACGAGATTTATGACACACATATATATATATATATATATATAT 867
DB 240 AGCAACTGAT 296
OY 868 TCCCTACACAAATTTACGACGACGACGACGACGACGACGACGACGACGACGACG 927
DB 297 AACTTCAAAACGATATGAGAGATATCAAAAGATATGATATATATATATATATATAT 356
OY 928 AATATCTCAAT 987
DB 357 GACTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1044
OY 988 TGAATATGACGAAAGATTTTAAATATATATATATATATATATATATATATATATAT

Db 2285 ATGCACAT-----AATTAAATTCAAATCTTAACTCCACACAAATTAATTAATTA 2232
 QY 1241 AAAAATTAGTAATTAATGCTGTGCTACTGTAATTAAGGATGAATGAATTA 1300
 Db 2231 CATATAAATTTAAATCTTACTCTCAAAAACCTAATTTAAATTAATTAATTA 2172
 QY 1301 ACACATATATAGCAAAAATTCATCAATAGATCTTCAATTAATTTGCTCATGT 1360
 Db 2171 ACACCTAATAAATTAATTAATAAATAAATTCCTTAAACCTTAAATTAATTA 2112
 QY 1361 CTGGAGATCTAATCTAATGTAACAGACAGACATTCAGCAATATGCGTTTGA 1420
 Db 2111 TATATAAATAATTAATCCGTTAATCAATTAATTAATAAATTAATAAATA 2052
 QY 1421 GAAATAGAAAAGTAGAATAATTAATTAATTAATTAATTAATTAATTAATTA 1480
 Db 2051 AAACAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1394
 QY 1481 GGAACAGATCTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1540
 Db 1393 AAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1334
 QY 1541 ATTGCAATATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1600
 Db 1933 ATTA---ACACACACACAGAAAACCAATTAATTAATTAATTAATTAATTA 1878
 QY 1601 AACATATATCTGTAATCTATCATCTTGTGTAATTTCAACACATGTAATTAAT 1660
 Db 1877 ACCCTAATTAACCAATCAAAATCTCTCCAAATGCACTAAA--AACTTTAA 1820
 QY 1661 AAAAATATAGAAATCAAGTTCACACATATTAATAAAGCAATCAATCCATCT 1720
 Db 1819 AAAAAAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1760
 QY 1721 CATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1780
 Db 1759 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1700
 QY 1781 GCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1836
 Db 1699 AATCAATGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1640
 QY 1837 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1896
 Db 1639 TCACCATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1580
 QY 1897 TCACCATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1956
 Db 1579 TTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1520
 QY 1957 ATACATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2016
 Db 1519 CTAATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1460
 QY 2017 CATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2076
 Db 1459 CTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1400
 QY 2077 TATGCCCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2136
 Db 1399 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1340
 QY 2137 TGCAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2196
 Db 1339 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1280
 QY 2197 CCACCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2256
 Db 1279 CCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1220
 QY 2257 CATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2316
 Db 1219 CTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1160

QY 2317 CTTCACACAAACAACTCAAGTTCATTAATTAACAAATTCAGAAATTAATTAATTA 2376
 Db 1159 ATCAAAATCCCTCCCTCCCAATCTACAAACCCAAAAAATAAATAATTAATTA 1100
 QY 2377 AACACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2436
 Db 1099 ACCCCAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1042
 QY 2437 TTCATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2496
 Db 1041 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 982
 QY 2497 TTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2501
 Db 981 TCTAA 977
 RESULT 11
 US-10-204-708-6/c
 ; Sequence 6, Application US/10204708
 ; Patent No. 6677731
 ; GENERAL INFORMATION:
 ; APPLICANT: OLEK, Alexander
 ; APPLICANT: PIEPENBERG, Christian
 ; APPLICANT: BERLIN, Kurt
 ; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
 ; TITLE OF INVENTION: By Assessing DNA Methylation
 ; FILE REFERENCE: 5013.1012
 ; CURRENT APPLICATION NUMBER: US/10/204,708
 ; PRIOR FILING DATE: 2003-05-06
 ; PRIOR APPLICATION NUMBER: PCT/EP01/03971
 ; PRIOR FILING DATE: 2001-04-06
 ; PRIOR APPLICATION NUMBER: DE 10019058.8
 ; PRIOR FILING DATE: 2000-04-06
 ; PRIOR APPLICATION NUMBER: DE 10019173.8
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: DE 10032529.7
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: DE 10043826.1
 ; PRIOR FILING DATE: 2000-09-01
 ; NUMBER OF SEQ ID NOS: 98
 ; SEQ ID NO 6
 ; LENGTH: 6669
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
 US-10-204-708-6
 Query Match 2.2%; Score 106; DB 4; Length 6669;
 Best Local Similarity 44.0%; Pred. No. 5.9e-11;
 Matches 910; Conservative 0; Mismatches 1155; Indels 31; Gaps 10;

QY 680 ATTGCACTGTCAGACATGCTATATGATTAAGTGAAGGAGATCATAGAAATAT 739
 DB 6369 AAAACATCTCTCTCCACATATTTATATATCTTAAAAAATTAATTAATTTATCTT 6310
 QY 740 ACACATATAAACAATTACCGAAAAATTTAAATATTTACCAATATCATCTACAGAC 799
 DB 6309 AAACATACAAATATCTTAAAAACCGATCTTATATATCTTATATTAATTAATTCAC 6250
 QY 800 AACATATACGACCGACATATAGACGAATTTATCGACACATATCATATATATC 859
 DB 6249 AA-ATACAAAAAATAATTTCTATCTATTAATTAATTAATTAATTAATTAATTA 6191
 QY 860 TTACTATTCCTACACACATTTACGACCATATGACAAATACACAAAGAGAT 919
 DB 6190 TATACACAAATATCTTATATACAAAAAATAATTAATTAATTAATTAATTAATTA 6131
 QY 920 ATGTGTAATAATCTCAATATTTATTTAGTAAAGATGGAATTTACATCTCA 979
 DB 6130 ATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6071
 QY 980 TTCCAGTGAATATGACAGAAATTTTACATATATCAATTTTATCTCTGAT 1039
 DB 6070 ATACAAATTAATCTAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6011
 QY 1040 CCAAGTTTGTAAAGAAATGCAAAATCTCTCTGATATGATATGATCTGAT 1099
 DB 6010 ACATATATATTTAAAAAATAATCCATTTAAATTAATTAATTAATTAATTAATTA 5951
 QY 1100 AATAAAGACTTGTATATGATGATGAACAAATAGATGATGATGATGATGATGAT 1159
 DB 5950 AATTAATAATTAACAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 5891
 QY 1160 TACCATCAAGATATTA-TAGGTGACAGATGAGAGTCAATTAAGATATGCAAT 1238
 DB 5890 ATAAATACATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5831
 QY 1219 TGATGTGAATATGCTGGAATAATTTAGTAAATATGCTGTTGCTGCTGATTAAC 1278
 DB 5830 AATATTTATTAATTAATCT-ATATTTAAATTAATTAATTAATTAATTAATTAATTA 5776
 QY 1279 TAAAGCATGAATGAATTAACATATATATAGAAAAATCATTCATATGATCTTCA 1338
 DB 5775 CTCTTAACCTATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5716
 QY 1339 TACATTAATTTGCTCATGCTTGGAGATCTAATATTAATTAATTAATTAATTAATTA 1398
 DB 5715 CCTTCTATGCTCTCTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5659
 QY 1399 AGCAGATATGCTTTGACATGAATATAGAAAAAGTGAATATTAATTAATTAATTAATTA 1458
 DB 5658 CCCAACTCTCTATGACCTATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5599
 QY 1459 TCAGAACTTAAGCTTTGGAAGAAACGATCTGTAATAAACAATGATATTAAGCAA 1518
 DB 5598 AAACCTA-AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 5541
 QY 1519 AAACTTAATGATGAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1578
 DB 5540 AAAATCTCTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5481
 QY 1579 ATCACTAGATTAAGCATATCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1638
 DB 5480 ACCAAGCTCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5421
 QY 1639 ACAGATATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1698
 DB 5420 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5361
 QY 1699 GCAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1758
 DB 5360 CCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5309

QY 1759 AAGTATACCTTAATAATTTCTAGCTAATTAATTAATTAATTAATTAATTAATTAATTA 1818
 DB 5308 AATTTCTATCTTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5249
 QY 1819 TAAACATTTGTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1878
 DB 5248 AACCTTCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5189
 QY 1879 ATATGTTGACACTTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1938
 DB 5188 ATAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5129
 QY 1939 AACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1998
 DB 5128 AATTCAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5069
 QY 1999 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2058
 DB 5068 AATCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5009
 QY 2059 TATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2118
 DB 5008 CTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4953
 QY 2119 GGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2178
 DB 4952 AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4893
 QY 2179 AACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2238
 DB 4892 ACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4835
 QY 2239 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2298
 DB 4834 ACCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4775
 QY 2299 CAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2358
 DB 4774 TTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4718
 QY 2359 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2418
 DB 4717 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4658
 QY 2419 AAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2478
 DB 4657 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4600
 QY 2479 ATTACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2494
 DB 4599 ACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4584

RESULT 12
 US-09-328-352-333
 ; Sequence 333, Application US/09328352
 ; Patent No. 6562958
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary L. Breton et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
 ; FILE REFERENCE: GTC99-03PA
 ; CURRENT APPLICATION NUMBER: US/09/328,352
 ; CURRENT FILING DATE: 1999-06-04
 ; NUMBER OF SEQ ID NOS: 8252
 ; SEQ ID NO 333
 ; LENGTH: 1077
 ; TYPE: DNA
 ; ORGANISM: Acinetobacter baumannii
 ; US-09-328-352-333
 Query Match 2.2%; Score 105.8; DB 4; Length 1077;
 Best Local Similarity 47.1%; Pred. No. 4.7e-11;

Matches 440; Conservative 0; Mismatches 477; Indels 18; Gaps 3;

Qy 638 ATCGATATTAGTGCAGATGATCTTTGGCAGACATAGCATATGCACTGTACGACAT 697
Db 70 ATTACCGTTTAAAGGGGGGAGCTTTGGTACAGCTTATGGCAATTTAGTGTCTGTAAT 129
Qy 698 GGTATATGATTAATCTTATGGGAGCTGATCATGAAATATACATATATAACACTTAC 757
Db 130 GGGCGCATACCATGATCTGATTCGGATTCGGATTCGGATTCGGATTCGGATTCGGATTC 189
Qy 758 CCAAAAAATTTAAATATTTTACCAATATCATCTTACCAAGCAATATATGCAACGAC 817
Db 190 CATATTTATTAACCTTTATTTACCAAGCTTCACTTTAAATCTTCAATACGTCGATTTCT 249
Qy 818 AATATGACGAGATATATCTGACACAAATACATGATATTTATCTTATCTTCTTACCA 877
Db 250 GATCTTGACGAGGAGGATATGATGACAGATATTTATTTAGTGAATTTCCAGGAC 306
Qy 878 CAATTAGCACCATATGTACCAAAATACACAAACAGATATGTGTAATAATCTCA 937
Db 307 TCATTCCTGATGTTTAAACCAATAGC-----GCCATATATCTGACCAAGCT 357
Qy 938 ATATTAATTTGATGAAGATGATGAATATACATCACTCAATTTCCAGTGAATAGCA 997
Db 358 GTGATCTTTTACCAAGGATGATGAAGATTAACATTTAGCTTCAATGATGATATAT 417
Qy 998 GAAGAAATTTTACATATATATCAATTTTATCTCTGTGTCAGATTTGCTTAAGAA 1057
Db 418 CGTAGAAGATATACAGAAAGTCCGTATGCGATTTGTCAGAGCAATCTCGCTAAAGAA 477
Qy 1098 ATTGACAGACATCTCTGATGATATGATCTGCTGCTGATATATAAGAACTTGGGA 1117
Db 478 ATTATGAGAGATATGCTCTGATCTGATATGCAAGTATGCTGATCTGCTGAT 537
Qy 1118 TCATGATGAAACAAATAGATATGATCTTAAATAATATATACATCAAGATATATA 1177
Db 538 GCCGACACATGATCTCATATGATGCTTATTCGCGCTTTTGGTAGCATATGTCAT 597
Qy 1178 GGTGTACGATTTGAGCTGCTTAAAGACATATTTGCAATTTGATGTAATTCGCT 1237
Db 598 GGAATGATGATGAGGAGGCTTAAAGATATTTATTCGCTGCTATGGAATTTGAGCA 657
Qy 1238 GAAAAAATTTAGTAAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1297
Db 658 GCATATTAATTTGATGAATATATCAAGATATATTTTAAAGGCTGCTGCTGCTGCT 717
Qy 1298 AAAACATATATATAGCAAAAAATCATCTTCAATATGATCTTCAATATTTGCTCAT 1357
Db 718 A-----GCCGTTTCGGGATTAACAAAGTGAACCAACCTTCAATGCTTATCT 771
Qy 1358 TGTCTGAGATCTATATATTTATATGATGATGATGATGATGATGATGATGATGATG 1417
Db 772 GGAATGATGATATTTTATGATGATGATGATGATGATGATGATGATGATGATGATG 831
Qy 1418 CTAGAAATGAAAAAGTGAATATATATATATATATATATATATATATATATATAT 1477
Db 832 TATCGCTAGCTGCTGATTAACCAAGCTGATCAAGCTGATTAAGCTGATTAAGCTG 891
Qy 1478 GAAGAAACAGTATCTGTAACCAAGCTGATCAAGCTGATTAAGCTGATTAAGCTG 1537
Db 892 GAAGATATATACAGATTTGATGATGATGATGATGATGATGATGATGATGATGATG 951
Qy 1538 CCAATTTGATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1572
Db 952 CCAATTTGATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 986

RESULT 13
US-10-204-708-23/c
; Sequence 23, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander

APPLICANT: PIERENBROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
FILE REFERENCE: 5013.1012
CURRENT APPLICATION NUMBER: US/10/204.708
PRIOR FILING DATE: 2003-05-06
PRIOR APPLICATION NUMBER: PCT/EP01/03971
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: DE 10019056.8
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: DE 10019173.8
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 98
SEQ ID NO 23
LENGTH: 11049
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-23

Query Match 2.2%; Score 104.2; DB 4; Length 11049;
Best Local Similarity 47.2%; Pred. No. 1.4e-10;
Matches 451; Conservative 0; Mismatches 493; Indels 11; Gaps 4;

Qy 1507 ATCATAGCAAAAAAATCTTAATGATGATGATGATGATGATGATGATGATGATGATG 1566
Db 7866 AACTAT 7807
Qy 1567 ACATGAGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1626
Db 7806 AAAAAATACGATCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7747
Qy 1627 TTGATTAAT---TTCAACATGATGATGATGATGATGATGATGATGATGATGATG 1683
Db 7746 AAT 7687
Qy 1684 CCACAT 1743
Db 7686 AAT 7627
Qy 1744 TTACATTAACAACTTAATGATGATGATGATGATGATGATGATGATGATGATGATG 1803
Db 7626 TAAAAAATCGCAACAAATTAACAAATTAACAAATTAACAAATTAACAAATTAACAA 7567
Qy 1804 ATCTGCTATCTCAATTAACATGATGATGATGATGATGATGATGATGATGATGATG 1863
Db 7566 AAAAAAATTTCTATATTAATCTTAATTTACAAATTAATTAATTAATTAATTAATTA 7510
Qy 1864 CTTATTTAT 1923
Db 7509 ATTACAAACAAAT 7450
Qy 1924 AATTTAT 1983
Db 7449 AATTAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 7390
Qy 1984 TATTTAGTACGAT 2043
Db 7389 AATTAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 7330
Qy 2044 ATTCAACATTAAT 2103
Db 7329 CCTTACAAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7270
Qy 2104 CTAATTAACCAAAAGGTTAAATATATATATATATATATATATATATATATATATAT 2163
Db 7269 TCCATAAAAAAT 7210

QY 2164 AGCTCAATTAATTAACCAATTTCTTAATTAACCACTTACTTAACATAGAACCAT 2223
DB 7209 AGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 7150
QY 2224 AGTATATCACTAGTATAGATATCCCTACAGCATGTATCAATTAATTAATTAATTCG 2283
DB 7149 TAAATTTAAAAAACAACCAAAACAA---CAAACAAAAAACAACCAAAACCTTAATTA 7094
QY 2284 TTAATTAATTTTAAACAATTAACCAATTTTCCAACTTCAACAACAACCAACTTCTAT 2343
DB 7093 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 7034
QY 2344 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2403
DB 7033 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6975
QY 2404 CATCATATCTTTTAAAAAGATGACTTCAACATTCATACATTAATTAATTAATTA 2458
DB 6974 AATCATATCTTTTAAACGAAACGATACATCTATATCCCAACACTTTAAAA 6920

RESULT 14
US-10-204-708-73/c
; Sequence 73, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEX, Alexander
; APPLICANT: RIEBENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; PRIOR FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO: 73
; LENGTH: 5152
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-73

Query Match 2.2%; Score 104; DB 4; Length 5152;
Best Local Similarity 43.9%; Pred. No. 1.3e-10;
Matches 987; Conservative 0; Mismatches 1220; Indels 41; Gaps 11;

QY 273 TTTTACATTAAGAAACAACCTGATTAACAAAGAACTCTCTCTGTTCCATCAATTTT 332
DB 4735 TTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4676
QY 333 CCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 391
DB 4675 ACAATTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4616
QY 392 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 451
DB 4615 ACTTTTAAAAACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4556
QY 452 CAAATTAATTAAGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 511
DB 4555 TCACATTAAT--ACTTTAAAAATCTTAATTAATTAATTAATTAATTAATTAATTAATTA 4498

QY 512 ATTTAGCATATATATCATCTGACATGATTAACAATAGTAACATTTCTTAACCTAGA 571
DB 4497 AAACATCAACATATTAATTTTAAACCTTTTAACTTTTAACTTTTAACTTTTAACTTTTAA 4438
QY 572 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 631
DB 4437 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4378
QY 632 ATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 691
DB 4377 ATACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4318
QY 692 GCACATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 751
DB 4317 CCGAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4258
QY 752 ACTTACGAAAAAATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 811
DB 4257 ACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4198
QY 812 ACCGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 871
DB 4197 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4138
QY 872 ACACAACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 931
DB 4137 CTACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4079
QY 932 ACTCCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 991
DB 4078 AACTCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4019
QY 992 ATACGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1051
DB 4018 ATCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3959
QY 1052 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1111
DB 3958 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3899
QY 1112 GGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1171
DB 3898 TTAATTTTA-----AAATTTAACTTAATTAATTAATTAATTAATTAATTAATTAATTA 3846
QY 1172 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1231
DB 3845 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3786
QY 1232 ATGCTGAAAAAATTTAGTA--ATAATGCTGTTGCTGCTGTTAACTTAAGCATGA 1289
DB 3785 TCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3726
QY 1290 ATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1349
DB 3725 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3666
QY 1358 GTCCATCATGCTGTGAGATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1409
DB 3665 TTAACAACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3606
QY 1410 CT-TTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1468
DB 3605 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3546
QY 1469 AAGCTTTGAAGAAACGATGCTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1528
DB 3545 AAAAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3486
QY 1529 GTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1588
DB 3485 TTA-ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3427

QY 1589 AAGCCATATCAACATATTTCTAGTCTATCATCTCTTGATTAATTTGACACATATGAT 1648
 DB 3426 TAAACATATTTATATTAACCAAAATTAATTTTAAAA-----AAATATTTACACCAAAATAT 3373
 QY 1649 AAATTAACGGTCAAAATATAGATACACGCTTCAACACATATATATAAAGCAGTCAAT 1708
 DB 3372 TTTAAAAATATTAACATTAATTTTCAAAACATTAATCATTAATACCTTTAAAAATTTAT 3313
 QY 1709 ACCCATGTTAATCAATCCAACTAGTATGTAATTAATTAATTAATTAATTAATTAAT 1768
 DB 3312 TTTAAAAAACTATTCACCAACATTTTATTAATTAATTAATTAATTAATTAATTAATTAAT 3253
 QY 1769 TAAATTTCTAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1828
 DB 3252 TATCTTCAATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3193
 QY 1829 TTAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1888
 DB 3192 AAATTTCTCAAAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3133
 QY 1889 AATCTTTTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1948
 DB 3132 CTAAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3073
 QY 1949 AAGAACCAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1999
 DB 3072 CATTAACATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3013
 QY 2000 TTACATGATGCTCTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2059
 DB 3012 TAAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2953
 QY 2060 ATGATTAATCTGTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2113
 DB 2952 ATTATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2893
 QY 2114 AAAAAGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2173
 DB 2892 TTAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2833
 QY 2174 AATTAACCAATTTCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2233
 DB 2832 ACCTCTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2173
 QY 2234 ACCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2293
 DB 2172 AATAAACTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2113
 QY 2294 TTAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2353
 DB 2112 CAATCCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2658
 QY 2354 TCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2413
 DB 2657 TATAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2598
 QY 2414 TTTTAAAAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2473
 DB 2597 TATTTAAACCAAAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2538
 QY 2474 GAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2501
 DB 2537 TTAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2510

RESULT 15
 US-10-204-708-14/c
 ; Sequence 14, Application US/10204708
 ; Patent No. 6677731
 ; GENERAL INFORMATION:
 ; APPLICANT: OLEK, Alexander
 ; APPLICANT: PIENBROCK, Christian
 ; APPLICANT: BERLIN, Kurt

; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
 ; FILE OF INVENTION: by Assessing DNA Methylation
 ; FILE REFERENCE: 5013.1012
 ; CURRENT APPLICATION NUMBER: US/10/204.708
 ; PRIOR FILING DATE: 2003-05-06
 ; PRIOR APPLICATION NUMBER: PCT/EP01/03971
 ; PRIOR FILING DATE: 2001-04-06
 ; PRIOR APPLICATION NUMBER: DE 10019058.8
 ; PRIOR FILING DATE: 2000-04-06
 ; PRIOR APPLICATION NUMBER: DE 10019173.8
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: DE 10032529.7
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: DE 10043826.1
 ; PRIOR FILING DATE: 2000-09-01
 ; NUMBER OF SEQ ID NOS: 98
 ; SEQ ID NO: 14
 ; LENGTH: 6113
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
 ; US-10-204-708-14
 Query Match 2.2%; Score 102.2; DB 4; Length 6113;
 Best Local Similarity 43.9%; Pred. No. 2.9e-10;
 Matches 948; Conservative 0; Mismatches 1163; Indels 48; Gaps 10;
 QY 362 TATTCGATCATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 421
 DB 6062 TATCCAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6003
 QY 422 AAATCTCAATAGCTTTTACCAAGAACTACAAATTAATTAATTAATTAATTAATTAATTAAT 481
 DB 6002 AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5943
 QY 482 ATATTTCAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 539
 DB 5942 AAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5883
 QY 540 TTAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 599
 DB 5882 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5823
 QY 600 TGACGTGTTTCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 659
 DB 5822 TTAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5768
 QY 660 CATTGGCAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 719
 DB 5767 TAAATCCATTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5708
 QY 720 GAGGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 779
 DB 5707 CTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5648
 QY 780 CCAGATCATCTTACCAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 839
 DB 5647 AACCAACATCTACCTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5586
 QY 840 ACACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 899
 DB 5587 AAACCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5531
 QY 900 AAATCAACACAAAGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 959
 DB 5530 AACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5471
 QY 960 TCGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1019
 DB 5470 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5417
 QY 1020 CAATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1079

Db 5416 AAAACCTTACTTAATATAAAAAACCAATCCAAAAATTCATATATATATATATAC 5357
 1080 GATATGCTCTGCTGTGATATATAAGAACTGGGATGATCATGTGATATAACATTAAGTA 1139
 5356 ATATATAACATCTCAATATAAAAAATATATATACAAAAACCAATCAATTAATCTA 5297
 1140 ATGATGCTTAAAAATATATATACATCAAGATATATATAGTGTGATGAGCTGAT 1199
 5296 AATTTAAAAATATAAAAAATTAATACAAAAATTAATCAATATAAAAAATATATCAAAA 5237
 1200 TAAAGAACATATGCAATGATGATGATATGCTGGAATAAATTTAGATATATG 1259
 5236 TAAAGAACATATGCAATGATGATGATATGCTGGAATAAATTTAGATATATG 5177
 1260 CTGTTGCTACTGTATTAATTAAGGATGATGATGATTAACATATATATAGCAAAA 1319
 5176 TTTATATAACT-----ATACCAATTTTAAAAATTTTCTATATTTTAAAAATA 5123
 1320 ATCAATCAATGATCTTCACTATATATGTCATCATGCTTGAGATCTATATATTA 1379
 5122 AATCTAAAAAAACCACTAATTAACAAACAAAAAACTAAAAAAATTAATATACAC 5063
 1380 CATGTCAACAGAACATGACGATATGCTTTGAGCTGATGATGATGATGATGATGAT 1439
 5062 AACACCTACTATATAAAAACTAAAAATTAAAAAATCAATCAATATATATATATAA 5003
 1440 ATATATAATCAATTAATGATCAACCTTAAGCTTGTGAAGAACAGTACTGTAAC 1499
 5002 TATATAACCACTAAAAATCTCATACCTCTATATAAAAATTAACCAATCAACCTT 4943
 1500 CACTGATCATATAGCAAAAAAACTAATGTAAGTCAATTTGATGATATATATTA 1559
 4942 AATCAATTTCTTAAAAATTAATCAATCACTCAACAAATTAATCAATCAATTTCT 4883
 1560 AATTTATCATGAGAT 1619
 4882 AATATATCATGATAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4823
 1620 TCATCTTGTATATTTCAACCAATGATATATATATATATATATATATATATAT 1679
 4822 CATATATTTAT 4763
 1680 TTCACCAAT 1733
 4762 TAAAT 4703
 1734 ATGAT 1793
 4702 AATCTAT 4643
 1794 ACAAATTCCTATCTCTATCTCAATTAACCAATGATATATATATATATATATAT 1853
 4642 CCAAAAAACAAAAAAATTAATATATATATATATATATATATATATATATATAT 4583
 1854 TAAATCTTTAC--TATATATATATATATATATATATATATATATATATATAT 1910
 4582 TCTAACCAACGACATCTCATCTATATATATATATATATATATATATATATAT 4523
 1911 AT 1970
 4522 AATATCAAAACCAAAATTTCAAAACCAATCAATCAATCAATCAATCAATCAATCA 4463
 1971 TTAAT 2025
 4462 TAAAT 4403
 2026 --TTTAT 2083
 4402 CGAAT 4343
 2084 ATTAGTCTTAT 2143

Db 4342 ACATCCAACTTAAACGACAAAAAGAACTCCATCTCAAAATATAACTAAATTAATTA 4283
 2144 ATATGAAGAGATATCAATCACTGATCAATTAATTAATTAATTAATTAATTAATTA 2203
 4282 AAT 4223
 2204 CTCTAAACATGAGAACCAATATATATATATATATATATATATATATATATATAT 2263
 4222 TCCAT 4167
 2264 CAAT 2323
 4166 AAT 4107
 2324 AACAACTCAAGTTCTATATATATATATATATATATATATATATATATATATAT 2383
 4106 CAATCTCAAT 4053
 2384 AGTTATCAATTTCTAT 2443
 4052 ACTTCAAT 3993
 2444 AT 2502
 3992 AAT 3934

RESULT 16
 US-09-328-352-1302/c
 ; Sequence 1302, Application US/09328352
 ; Patent No. 6562958
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary L. Breton et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
 ; FILE REFERENCE: GTC99-03PA
 ; CURRENT APPLICATION NUMBER: US/09/328,352
 ; CURRENT FILING DATE: 1999-06-04
 ; NUMBER OF SEQ ID NOS: 8252
 ; SEQ ID NO 1302
 ; LENGTH: 1434
 ; TYPE: DNA
 ; ORGANISM: Acinetobacter baumannii
 ; US-09-328-352-1302

Query Match 2.1%; Score 101.4; DE 4; Length 1434;
 Best Local Similarity 45.1%; Pred. No. 3.2e-10;
 Matches 516; Conservative 0; Mismatches 611; Indels 16; Gaps 3;

Db 2543 GCAATTCGATATCACTTACTCTCAAAACGATATATATATATATATATATATAT 2602
 1313 GCAATTCGAT 1254
 2603 AATTCCTCAAGACTTTACGTAATGTAAG-----AATTAACCTTACCTTAAACAA 2656
 1253 AATTTTTCAGACTTGAGAGAGCTGCACAGTGTGTATGCTTTGAGACGAGCAAGA 1194
 2657 TCATTAATGAGCAATTTAGTATATATATATATATATATATATATATATATATAT 2716
 1193 AAGGATATGAAATATATGCTCTTGAAGAGAGAGATGTAATGATATCATCTATCGTA 1134
 2717 TTGAGATTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2776
 1133 GCATCTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1074
 2777 GGCATTCAT 2836
 1073 GGAAGGTTGCTGATGAGATATATATATATATATATATATATATATATATATATAT 1014
 2837 CCAATTTTTCAGATGCTAT 2896
 1013 TCCCAATCCATCAATTTTTCATATATATATATATATATATATATATATATATATAT 954

Tue Apr 6 07:57:22 2004

us-10-081-051-52.rni

Page 16

Db	3200	AAATTCCTAATTAATATAAAAAAATATAATTTATTTAAAAACCGCAATTTATGCTTATAA	3141
Qy	2366	TATACTTCACCAACATATGATTCACAAATTTCTATAATACATCATATCCCTTTTAAAAAGA	2423
Db	3140	TCCTCCCCCAAAATATCATCAAAAAATTACTTTTATTAACATCAAAAAACAATCTTAAATTT	3081
Qy	2426	TGACTTCACAAATTCATACATTAATATATATACATGCA---AGTTATTCAGAAACATTA	2482
Db	3080	TTCTTCACAACTACTACTACATTAATAAAAATATATATACATATATATATTTTACATPACA	3021
Qy	2483	CACAAATCCAGAAATTTTCAATTTACAGGTGAATGCTGAACGTTCATTCACAAATTTCA	2542
Db	3020	ATATATAAAAATATATTAACATATACATTAATTAATATATATATAAAAACATTAATATCCA	2961
Qy	2543	GCAGTTTCGTATACATTAACCTCCAAAACGA-----GATMACTGAGTTTTTAAATCTG	2595
Db	2960	AAATTTAACTTAAAAAAAATATAAAAAAACTTTTGTTACTATATATTTTAAATACAT	2901
Qy	2596	TCCTTCAATTTCTTCACAGCTTTACGTAATGTAACGAATTAACCTTTAGCACTTAAAC	2655
Db	2900	AAATTTAAAAATTTTAAATCTTTAAACCTATTTCCAAAATATTTTATTTTAAATCTTTAAT	2841
Qy	2656	ATCATTAATAGCGGAATTAGATACTATATCTACGTGTAATCTTTGCTGATCATCTC	2715
Db	2840	TTTCATCAAAATATCATATATATATATTAATTTACTACGCAAAAATCTCCATTAACATTTAA	2785
Qy	2716	CTTAGAGATTTCTATATTTAAAAATCCATCTCTATACATTAACGTATATGCTTAACTTAC	2775
Db	2780	CACAAAAATATTCACACATTAACCTTAACCTTTAAAAATTAATCTATATATCAATTCACACCC	2721
Qy	2776	TGCGCATTCATATGACTGCAATTCATTAATTAAGCTTCATCATTAATATACATAGATACAAA	2835
Db	2720	ACAACCTTAACATATATCATATCATATCTTAATPAA--TAAATPAAACAAATATACCTCAATACA	2663
Qy	2836	ACCTATTTTTCACAGATCTATTCATATAATATACCTACAAATTTCCGGTATATCTGTACA	2895
Db	2662	TACAAATCTCTTCACATTAATAAAAACAAACCTATATATACAAATATTAATATTTAAAA	2603
Qy	2896	ATATTCCTACTAAAGACGGTACTCTAATATGAGAGACATTTAATCTATATATTAATCTC	2955
Db	2602	ACACTATATCTTAATATAAAAAAAATCTAACATCATATATATATAAAAAAATCAATATGCC	2543
Qy	2956	ACAAATCCCTACAGCTTTTACTCTCATCTTCAATATCTTTGGAGAGAGACAAATATTTCT	3013
Db	2542	TAAATTAACCAAAATTTAAAAAAATATAAAAAAATATTAATTAATTAATTAATTAATTAATCT	2483
Qy	3016	CACATCTATACTACACGAAATCTTACTATT	3045
Db	2482	TTTACGATATCTAAAAATATCTTATATCTT	2453

RESULT 18

US-10-204-708-60/c

Sequence 60, Application US/10204708

Patent No. 6677731

GENERAL INFORMATION:

APPLICANT: OLEK, Alexander

APPLICANT: PIEPERBROCK, Christian

APPLICANT: BERLIN, Kurt

TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication

TITLE OR INVENTION: by Assessing DNA Methylation

FILE REFERENCE: 5013.1012

CURRENT APPLICATION NUMBER: US/10/204,708

CURRENT FILING DATE: 2003-05-06

PRIOR APPLICATION NUMBER: ECT/EP01/03971

PRIOR FILING DATE: 2001-04-06

PRIOR APPLICATION NUMBER: DE 10019058.8

PRIOR FILING DATE: 2000-04-06

PRIOR APPLICATION NUMBER: DE 10019173.8

PRIOR FILING DATE: 2000-04-07

PRIOR APPLICATION NUMBER: DE 10032529.7

PRIOR FILING DATE: 2000-06-30

[illegible]

D	b		5238	TAAACAACATTAACCATCAATTAATTCATAAATCAATGACATCTTCCATTAAACC	5179
O	y		1042	AAGTTTGCTAAAGAAATTCGAGA-CATCTTCTGTGGTAGTAGACTTGCCTGGATGA	1100
D	b		5178	AAAACTACTTTTTTAAACAACCTACATTAATTTTCAATTAACATTTTAAAATC	5119
O	y		1102	TAAAGACTTGGTGAATC--ATTGATGAACAATAAGATGATGTCTTAAAAATAT	1158
D	b		5118	CTTACTACTTTTAAATCTCAACACCCACCATATCTATTAACTTAATTAACATAT	5059
O	y		1159	ATACCATGAAGATTAATATAGGTGACATTTGGAGCTGCATTAAAGACATATTCAT	1218
D	b		5068	CTCTAACCTTCAATTAATTTCCATTAATCCAAATATCCACAAATTAATTAACATA	4999
O	y		1219	TGCATGTGAATTAATCGCTGGAAAAAATTAAGTAAATATGCTGTGCTACTGTATAC	1278
D	b		4998	AAATCCCTTCACGCTACTAAAAAAAATTTAAATTCATCTACAAACCAACATATATA	4939
O	y		1279	TAAAGCATGATGAATTAATAACCTATATAGCAAAAAATCATTAATGATCTCA	1338
D	b		4998	TATATCGAAATTACTAATTAATCTTATCTTAATTAACATTAACCAATTAATTAATTTT	4879
O	y		1339	TACATTAATGGTTCATCATGTCTTGAGATCTAATATTAACATGTACAGACATTC	1398
D	b		4878	TAAATCATACATCAAAATTTCTTAATACCTCACGTAAAAAATTTAAAAACAAATC	4819
O	y		1399	AACGCAATTCGGCTTTGGACTAGAAATGSAAGAAAGTGAANATTAATACATTAATGA	1458
D	b		4818	TAAAAAATAAATATGCTAATATTAATTTTAAATTAATCTTTAACATAT- -AAAAAATA	4762
O	y		1459	TCACAACCTAAAGCTTGTGAAGAACAGTACTGTAAACCATGATATCATTAGCAA	1518
D	b		4761	CTACATCTTAATTTAAATCCATACACACCACTCATCAATAAAAATTAATTTAAAT	4702
O	y		1519	AAAACCTATGAGAACACTACCAATTTGGATATCTTTTACATTTATTAACATGGAAT	1578
D	b		4701	TAAATCAATTTCTACTCTTAAATATTAATTAATTAACCTATATTAATTAATAAT	4642
O	y		1579	ATCACTAGATAAAGCATATCAACAATATTAATCTTAGTCAATCATCTTTGATATTTCA	1638
D	b		4641	AAAAACCAAAATAAAAAACAATTAATTTCAATCAAAAAAATAAATTAATATAT	4582
O	y		1639	AACATATGATTAATTAACGCTCAAAAAATATGATACACAGCTGCCACATATTAATA	1698
D	b		4581	TTCAATCAATTAATTAACAAATTAATTAATTAACAGATTAACACTATATCCATTTTAA	4522
O	y		1699	GCACTCAAAATACCCATCGTTAAACATCACTAGATATGTAACATTAATTAACAACTT	1758
D	b		4521	AAATTACTTAATATTAATAAAATTAATCTTAACACTTCCATAAAAAATTTTAAATCA	4462
O	y		1759	AAGTTAACTTAATAATTTCTAGCTAA-TTATATTAACAAATTCCTATCTCATCAT	1818
D	b		4461	ATAAATTAATATA---TCTTAATTAATTAACAAAAAACCTTAATAAAATACTAAT	4466
O	y		1819	TAAACCAATGTATACGATTAATTAATTAAGTCAATTAACCTTTTACTTAATTAATA	1878
D	b		4405	TTTCCCTTCAAAATATTTATTAACCAACTTCAAAATTAATTTTCTCAAACTAAAAAT	4346
O	y		1879	ATATGTTGCACACTTTTCACTACATCATGATATTAATGAACATTAATTTACATAAAT	1938
D	b		4345	TATATCTTAATTAATTTTAAAAAATAAATCCCAACCCCTCCAAATTAACACCT	4286
O	y		1939	AACATTAATAAAGAACCAATACATTAATTAATTAATTAATTAATTAATTAATTAATTA	1997
D	b		4285	ATAAATTAACAAATTAATTAATTAATAATTAATTAATTAATTAATTAATTAATTAATTA	4226
O	y		1998	TATTACATAGATGTCTGACATTAATTAATTTATCATCTTGACATATTTCAACATPAA	2057
D	b		4225	TACATCTAAATTAATTAATTAACAAATTAATTTTATATCAATTAATTAATTAACACAC	4166
O	y		2058	ATATGATTAACATGTTTTTTTATGCCCATTAATGCTTAATTAATAATTCGTAATTAACCA	2117
D	b		4165	TTCTAAAAAATATTAATAAAAAAAAAACATTAATAATTAACACCTTAATTCATTAATTAAT	4106

```

OY      2118  AGGTTAAATTCACATCTCCGCAACAAATATGAGAAGTACATACACGTCGAAATTAAT  2177
DB      4105  TAAATAAAAAACCTCTTACAACTCTTTAAATTTACATATAAAAACCTACACCAAAAATTC  4046
OY      2178  AAACAAATTTCTAATATATACCACTACTCTTAAACATAGGAACCAATGTATATACACT  2237
DB      4045  CATACCAAAACCATTAATATATAAACCTCCCACTCAATATATACATATAAAAATATCTTAA  3986
OY      2238  AAGTACAGTATCCCTACAGCATGTATCAAAATATATAAAGAAAAATCGTATGTATATAT  2297
DB      3985  CTTTAAAAAATCAATA-AAAATATATTAATACCTTAAATATAAAAATAAAAAATTTATTA  3927
OY      2298  ACAATTCAGAAATTTTCCAACTTCACACACAAACAACTCAAAAGTCTATTATTAACAATTC  2357
DB      3926  CTAACCTTAAATATACGATATTAAC----TCTCCCACTAATTATTAATATACCACTATTA  3871
OY      2358  GAATTAATATACCTCCACCAACATATGTTATCAAAATTCATATATACATATACCTCTT  2417
DB      3870  AATATTAATATCCACAA-ATTTTAAATATTAATAATCTATATATATTTCTTAAAAAATAT  3812
OY      2418  TAAAAAGTAGCTTCACAAATTCATACATATTAATATATA  2456
DB      3811  TATAACATTAATATTTTAAATTTAATTAATATATATAA  3773

RESULT 19
US-10-204-708-32/c
/ Sequence 32, Application US/10204708
/ Patent No. 6677731
/ GENERAL INFORMATION:
/ APPLICANT: OLEK, Alexander
/ APPLICANT: BERLIN, Kurt
/ APPLICANT: PIENBROCK, Christian
/ TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
/ TITLE OF INVENTION: by Assessing DNA Methylation
/ FILE REFERENCE: 5013.1012
/ CURRENT APPLICATION NUMBER: US/10/204, 708
/ PRIOR FILING DATE: 2003-05-06
/ PRIOR APPLICATION NUMBER: PCT/EP01/03971
/ PRIOR FILING DATE: 2001-04-06
/ PRIOR APPLICATION NUMBER: DE 10019058.8
/ PRIOR FILING DATE: 2000-04-06
/ PRIOR APPLICATION NUMBER: DE 10019173.8
/ PRIOR FILING DATE: 2000-04-07
/ PRIOR APPLICATION NUMBER: DE 10032529.7
/ PRIOR FILING DATE: 2000-06-30
/ PRIOR APPLICATION NUMBER: DE 10043826.1
/ PRIOR FILING DATE: 2000-09-01
/ NUMBER OF SEQ ID NOS: 98
/ SEQ ID NO 32
/ LENGTH: 8093
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-32

Query Match      2.1%; Score 101; DB 4; Length 8093;
Best Local Similarity 47.1%; Pred. No. 5.1e-10;
Matches 546; Conservative 0; Mismatches 595; Indels 18; Gaps 7;

OY      1512  TAGCAAAAAAAGCTTAATCTAGACACTACCACTTTCGATATCTATTTACATTTATTAAGT  1571
DB      6716  TAACTTAATTCCTTAAATTTACTAAATTAATATACCACAAAATTAATAAATTTTCAATTA  6657
OY      1572  AGAATATATCACTAGATTAAGGCATATCAAAACATATAT-CTTAGTCTATCATCTTTGA  1630
DB      6656  AAACAAAAGCTTAAACCAAAATTTTAAATTAATATATCTTAAAAAATCTATTAATA  6597
OY      1531  TAATTTCAACAGCATGATTAATTAACGCTCAAAAATATAGATCAACGTTACACACATA  1690
DB      6596  TAAATTAATTCCTTATATATTAATTTTACATATATAAAGCTTAAATTTACCTTTTATA  6537

```

1691 TATAAAGAGGAGCAATATCCATGCTTACATCCAACTAGTATGTAACATATTACATA 1750
6536 AATAACAAAATTTTAAATTTACTATTTTAAACAAAACAAAT---TATCTTAATATAAA 6480
1751 ACAAACTTAAGTAACTTAATTTCTAGCTAATTAATTAACAATAATCTCTCTCT 1810
6479 TACAAATTTACTACATCTTAAT---AAACTCTTTTAAATTTTCTACTTTATCT 6422
1811 ATCTCATTTACCAATTTGTTATCAGATTATTAATAGTCAATTTAACTTTACTTATTT 1870
6421 ATCTTCATATTAATTAATTAATGACATTTCCAAATTTAAATTAATCTCTTACAT 6352
1871 ATTAATTAATTTGTTGACACTTTTTCATCATCATGATATTAATGAACATAATTTTA 1930
6361 CTAAACAAATCTTAATCTATATCATTTTAAACAAATAATTTTCTTAAACAAATTTT 6302
1931 CATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1990
6301 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6242
1991 TACGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2050
6241 CTATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6182
2051 ACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2110
6181 CATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6124
2111 CCCAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2170
6123 CCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6064
2171 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2230
6063 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6004
2231 CATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2288
6003 CTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5944
2288 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2348
5943 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5884
2349 CAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2408
5883 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT -A 5830
2409 TATCTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2468
5829 CTCTCTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5770
2469 ATTCGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2538
5769 CATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5710
2529 TTCCAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2588
5709 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5650
2589 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2646
5649 CTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5550
2647 ACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2665
5589 ACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5571

RESULT 20
US-10-204-708-20/c
; Sequence 20, Application US/10204708

Patent No. 6677731
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPENROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
FILE REFERENCE: 5013.1012
CURRENT APPLICATION NUMBER: US/10/204,708
PRIOR FILING DATE: 2003-05-06
PRIOR APPLICATION NUMBER: PCT/EP01/03972
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: DE 10019056.8
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: DE 10019173.8
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 98
SEQ ID NO 20
LENGTH: 6866
TYPE: DNA
ORGANISM: Artificial Sequence
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-20

Query Match 2.1%; Score 98.8; DB 4; Length 6866;
Best Local Similarity 46.8%; Pred. No. 1.3e-09;
Matches 453; Conservative 0; Mismatches 502; Indels 13; Gaps 4;

1495 AAAAAACATGATATCTATGACAAAAAACTTAATAGTAGACATGATGATATCTAT 1554
6283 AACAACCTCAATTTCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6225
1555 TTCAATTTTATCATAGAGATATATCACTAGTAATTAATTAATTAATTAATTAATTAAT 1614
6224 ATACATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1615
1615 GTCTATCATCTTTGATTAATTTCAACACATGATTAATTAATTAATTAATTAATTAATTAAT 1674
6164 AACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6105
1675 CAACGTTCAACATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1731
6104 CAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6045
1732 GATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1791
6044 TTCACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5985
1792 AAACAATTCCTATCTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1851
5984 TTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5925
1852 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1911
5924 ATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5865
1912 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1971
5864 TTCTTTTGGATATTTT---CAAAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5807
1972 TACTAAACCACTATTTACTAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2031
5806 AAAAAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5747
2032 CATCTTGACATTTTCAAAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2091
5746 TTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5687

QY	2092	TTAATTAATAATTGCTAATTACCCAAAAGSTTAAATATACACATTCCTGCAACAATGTGAAG	2151
Db	5666	AAAAATTAATAAAAAATTAACCTAATAATCAAAAAGCTTAT-----AAACAAAAT	5634
QY	2152	AAGATACATACACGCTCAAAATAAATTAAACAAATTTCTAATATACACGCTACTTAAAC	2211
Db	5633	AAAAAAGTATATAAATAAATAAACAATAAATACAAAACAAAACCAATTAATAAAAAAATA	5574
QY	2212	ATGGAACCAATGATATATCATACTAAGTAAAGATTCCTTAGACATGTATCAATATAT	2271
Db	5573	ATAAAAAAATTTAAATCTTAAATAAATAATATTAATTAATCTCATTAATAAATATTTAAACA	5514
QY	2272	AAAGAAAATCGTATGATATTTTATACATATTCACAAATTTCCACTTCAACCAACAAAC	2331
Db	5513	AAAAAAACATCAAAAACCATAAAAATATMAAAAACAACTTTAAACTTAAAAATATTAAC	5455
QY	2332	TCAAAAGTCTATTAATACAAATTCAGAAATTAATATATCTTACCAACAACATAGTTATCA	2391
Db	5453	TCATACCTATATCTCAACACTTTAAAAAAACCTAAAAACAACMAATTAAGTTAAAAACAAA	5394
QY	2392	AATTCATATATATCATATATCTTTTAAAAAGATGATTCACATTCATATCATATATTA	2451
Db	5393	ATTCAAAACCAACCTTAACCAATACGAAMCCCAACTCTAATATATATCAAAATATTA	5334
QY	2452	TTATACAT	2459
Db	5333	CTAAACAT	5336

```

RESULT 21
US-10-204-708-80/c
; Sequence 80, Application US/10204708
; Patent No. 6677791
; GENERAL INFORMATION:
; APPLICANT: OER, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO. 80
; LENGTH: 8961
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; FEATURES:
; NAME/KEY: unsure
; LOCATION: (5096)
; OTHER INFORMATION: n is a or g or c or t
US-10-204-708-80

```

	Query Match	2.0%	Score 95.8	DB 4	Length 8961
	Best Local Similarity	44.0%	Pred. No. 4.7e-09		
	Matches 939	Conservative 0	Mismatches 1152	Indels 42	Gaps 11
QY	397	ATAATGATGCTATCTTAATTAATAAAAAATCCTCAATAGCTTTTACGACAGAACTTCAAT			456
Db	7239	AAAATATTACTCTTATACCAACAAAAAATATAAACATTATACCACTTAAAAATTTT			7180

QY	45	ATATAGACACTTAATTATTAATTAATTTCAATTTATTAATCTCATCAATTA	516
Db	7179	TTTTTTTTTAAAAATATBACCCCAACAAATATAAAAACCTTTAAATA	7120
QY	517	GCATATATCATCTGACATGATTAACAATAGTACACATTTCTTAACATGATATA	576
Db	7119	CTATTAACCTCAAAATATATAATAAAAATACCTAAAAATATAAACATTAAAAATA	7060
QY	577	TACATTAATATATATATATTTTGACTGGTTCTGATTAATTTGGTAAAAAGAA	636
Db	7059	AACCAAAATTAATTTCAAAAAATAAAAAAACAATACTAATATATATTAATCTTTAA	7000
QY	637	AATCAGATTTAGTGCAGATCATTTGGCAGACGACATAGCATTTGCCTGCAGACA	696
Db	6999	AAAAAAA---ACATTAACATATTTTAATTTCAATTTAAAAATCTAATAATA	6944
QY	697	TGTATATCGATTAATCTTAGGGAGCTGATCATGAATATTTACATATTAACCTTA	756
Db	6943	AACATTTCTTCCAAAAAAATCCCTTAATAAATAAAAAAACAATAAAAAATCTATATA	6894
QY	757	CCGAAAAATTTAAATATTTAC-CCACATATCTTACAGACACATATATGCAACA	815
Db	6883	CTATATCTATACCCATTAATCTTTAAATATCTTCCAACTCTTAAAAAACGAGA	6824
QY	816	GCAATATAGAGAGATTAATCTGACACACATACATGATATTTACTATTCCTACAC	875
Db	6823	TAAATCTTATTAATATATCTCTCTTAAACCA---ATAAATTTATTTCTATTAAC	6767
QY	876	AACATTAAGACATATGTACACAATPACACACAAAGCATGTGTAACATCTC	935
Db	6766	AAAAAATTAACAAATCTCAACAATCTCCCTTAACCATCTATCTGTAATACATA	6707
QY	936	CATATTAATTTGATGAAGGTATCCAAATTAATCATCAATTTCCAGTGAATAG	995
Db	6706	ATTACTTAATTAATTAATCTTTTAAAAAAACCCCTAATCTTCTTAATCTGCAT	6647
QY	996	CAGAGAAATTTTCAATTAATCCAAATTTTATATCTC-----TCTGCTCAAG	10444
Db	6646	AATCGAAACCGACTTAATTAATCTATACACAACAATTAATTAATTAATTAAT	6587
QY	1045	TTTTCGTAAGAAATTCGAGACATCTCTCTGATGATAGTACTTCTGCTGATATTA	1104
Db	6586	AATTAATPACAAAAACAACTTCACTCTTAATATATTAACAAAAACCTAATATC	6527
QY	1105	AGAACTGTGTAATCATGATAGAAACAATAGTAATGATGTTCTAAATAATA---TA	1161
Db	6586	ACATCGTAATTAATTCATCCCAAAAAATTAACCTAATAATTAATTAATTAATCCCTA	6467
QY	1162	CCATCAAGATTTATAGGTGACAGATTTGAGCGCATTAAGAACATTAATGCAATGC	1221
Db	6466	ACCTAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	6407
QY	1222	ATGTGAAATATCGCTGGAAAAATTTAGGTAATATGCTGTGCTACTGTTATTACTAA	1281
Db	6406	AATTAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	6347
QY	1282	AGGATGAATTAATTAACACATATATAGCAAAAAATCATCATAGATCTTCAATAC	1344
Db	6346	ATCAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	6287
QY	1342	ATTATTTGTCATCATGCTTGGAGATCTAATTAATTAACATGTACAAACAGAACATTCAG	1401
Db	6286	CTAAAAAATACCTAAATTTTAAATTAACCTAATAATTAATTAATTAATTAATTAATTA	6222
QY	1402	CAATTTGGCTTTGGACTAGAAATAGGAAAAAGTGAATATTAATCATTTAATAGATCA	1461
Db	6226	GAATAT--CTAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	6169
QY	1462	CAACCTTAAGCTTGTGAAGAACCATGATCTGTAAACCACTGATATCATTTAGCAAAAA	1521
Db	6168	TACCAAAATCTACTTTAACCCCAAAAAAAATTAATTAATTAATTAATTAATTAATTAATTA	6109
QY	1522	ACTTAATGTAGACCTCAATTTGCAATCTATTTTAACTTAATTAATTAATTAATTAATTAATTA	1581

Tue Apr 6 07:57:22 2004

us-10-081-051-52.rni

Page 20

Db	6108	AAAAAAGCTAAATPACCTACCTAACCTCCCTCCTACTCTAAATATTAAATTAACCTAACAA	6049
Qy	1582	ACTGATTAAGCCATATCAAACATATTAATTGATGCTATCATCTCTTGATATTTGACAA	1641
Db	6048	CGCCTTAAAAATPAAAAATPAAAAATATPAAATATATPAAATACCTAACCTGAAAAATAT	5989
Qy	1642	CAATGATTAATPAAAGCTCAAAAAATATAGAAATACAGCTTGACCCGATATPATAAAGCA	1701
Db	5988	AAAAATATATPAAAAATPAAAAATPAAAAATATATPAAATATATPAAATATPAAAAAT	5929
Qy	1702	GTCAAAATCCCATGCTTACATCCACCTPAGTAGTAAACATATTACATPACAACT-TAA	1763
Db	5928	TTATATATATATPAAATPAAACGAACTCCCTACATATTTATPAAATATPAAATATPAA	5868
Qy	1761	GTATPACTTAAATTCCTAGCTAATPATAAACAATTCCTATCTCTCAATTA	1823
Db	5808	TCCATCTPATAAAGCAACATPAAACTATPAACTATPAAATATATCTTTTAAATPAAAA	5749
Qy	1873	AAATPAAATATGTTGACAGCT--TTTGACATACATCATGTAATPATAACATATTTTA	1930
Db	5748	AAATACTTACTATPAACTATPAACTPAAAAATPAAAAATPAAAAATPAAATPAACTC	5689
Qy	1931	CATPAAATPACTAATPAAAAAGAACCAATPACATATPAAATPACTPAAAAACCTATTTAC	1990
Db	5688	TAAACCTPAACTACAAAACTPAAAAACACTCPAAAAATPAAATPAAATPAAATATCA	5629
Qy	1991	TACGTAATATPACATAGATGTCCTGACATPAAATTCCTATCAGCTTGACATATTTCAA	2056
Db	5628	ATATPAAAAATPAAAAATPAAATPAAATPAAATPAAATPAAATPAAATPAAATPAAAT	5688
Qy	2051	ACATPAAATATGATPAACTGTTTTTTTATGCCCCATPAGTCTTATPAAATPCTPAAATTA	2110
Db	5568	TATATATATPAAAAATPAAAAATPAAATTTCTAATCTATPAAATPAAATPAAATPAAAT	5509
Qy	2111	CCCAAAAAAGTTAAAAATCACATTCCTGACAAATA---TGAGAAATPACATPACACG	2166
Db	5508	TAAAAATPAAATATATPAAATPCTPAAAAATPAAATPAAATPAAATPAAATPAAATPAA	5449
Qy	2167	TCAATATATPAAACAAATTTCTAATPAAACACCTACTCPAAACATPAGAACCATATAGT	2228
Db	5448	ACATPAAATPACAGAAAAATCTTATPAAATPAAATPAAATPAAATPAAATPAAATPAA	5389
Qy	2227	ATATCATCTPAAAGTAAGTATCCTPACAGCATGATPAAATPAAATPAAATPAAATPAAAT	2288
Db	5388	TTATTTATPAAATATPAAATPAAATPAAATPAAATPAAATPAAATPAAATPAAATPAA	5329
Qy	2287	TGATATTTATPACATPACAAATTTCCAACTTCCACACACAAACCTCAAGTCTATTA-	2345
Db	5328	TTATPAAATPCTPAAAAATPAAATPAAATPAAATPAAATPAAATPAAATPAAATPAA	5269
Qy	2346	--TAAACAATPACAGAAATTTATPAAATPAAATPAAATPAAATPAAATPAAATPAA	2409
Db	5268	TCTCTPAAATPAAATPAAATPAAATPAAATPAAATPAAATPAAATPAAATPAAATPAA	5209
Qy	2404	CATCATATCTTTTAAAAAGTGACTTCAACATTCATACATPAAATPAAATPAAATPAAAT	2463
Db	5208	TATATATATATPAAATPAAATPAAATPAAATPAAATPAAATPAAATPAAATPAAATPAA	5149
Qy	2464	TGTTTATTCAGAAACATTPACAAATCCAAAGAT	2496
Db	5148	TAAAAATPAAACAAAAATPAAACGATTAACCAAT	5116

```

; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 3331
; LENGTH: 1023
; TYPE: DNA
; ORGANISM: Proteus mirabilis
; US-09-543-681A-3331

Query Match      2.0%   Score 95.6   DB 4   Length 1023:
      Local Similarity 47.3%   Pred. No. 3,66-09:
Matches 463; Conservative 0; Mismatches 499; Indels 21; Gaps 5;

```

QY	638	ATCGATATTATAGGTGCAAGATCAATTGGACACGCAATACCAATGCACTGCATGCACAT	697
Db	22	ATGACAGTTATCGGTGCCGGTTTCAATACGGCACCGCCTTAGGATTAACCTTAGCGCGTAAT	81
QY	698	GGTATATCAGTTAACTTATGGGGAACGGATCATAGAAATATTACACATATTAACACTTAC	757
Db	82	GGGACAGCATGTTGCTGTGGGGGACAGATCCAGACAGTGGGGCATTAAGAACAAAGC	141
QY	758	CGAAAAATTAAATATTTATACCACATATCATCTACCAACAATATATGCAACGAC	817
Db	142	CGCTGATATCAAGCCTTCTGCCGATGTTCCCTTCCTGATAGTTATATATAGGAAGCT	201
QY	818	AATATPACGAGTATTATCTGACACAAATACATGTATTTATCTTAACTATTTCTTACCA	877
Db	202	TCCTTGGCAAAAGGATTTGAAGCAGCGCTGA--ATATCTTGGGTGATCCCAAGCAT	258
QY	878	CAATTACGACCAATATGTACAAATATACACAAAGCATATGTGTAATAAATACTCA	937
Db	259	GCTGTTGGTGAAGTACGACAAAT-----CAACCCCTTTTACGTCAGATGCGCGT	312
QY	938	ATATTAAATTGTAGTAAAGGTATGA--AATTACATCACTCAAAATTTCCAGTGAATA	994
Db	313	GTTGTTGGGCGACAAAGGCTGTGAAGACATCTGCTGCTATTATCAAGATGTTGCC	372
QY	995	GCAGAAAGAAATTTACATATATATCCAAATTTTATCTCTGTGTCGAATTTGCTCAA	1054
Db	373	CGTAAAGATTATAGGTAAATGAAATCCCGCTGCGAGTATATCTGCGCCTTACTTTGTCAA	432
QY	1055	GAAATTCGAGAACTCTTCCTCTGTAGATATAGTACTTCTGGTGTATTAATAAGAACTTGGT	1114
Db	433	GAGCTTGGCGGGGTTTGCCCAACCGCATGGCACTGGCGCTCGACGATTAATCTCTTTTA	492
QY	1115	GAATCATTTGATAGAAACATATAGTATGATGT--TCTAAAAATATATATACCATCAAGT	1171
Db	493	GAGCAGTTACAAACGCTATTTCATTGTGGTAAAGTTCCGAGCTATATAAAAACCTGAT	552
QY	1172	ATTATAGGTGACAGATGTGAGCTGCAATTAAGAACAATAATGCAATTGCACTGGGATA	1231
Db	553	TTTATTCGTTGCACTCGGTGGGCTGTTAAAACTGATGTGTTATTTGGTGGGGATG	612
QY	1232	ATCGCTGAAAAAATTTAGGTATATATGCTGTGTCTACTGTATTAACTTAAAGCATGAT	1291
Db	613	TCTATGTTATAGGATTTGGCGCTAAATGCGGCTACCCACATGATTAACCGAGGCTTAGCC	672
QY	1292	GAAATTAACACATATATATATGCAAAAAATCATTCAMPATGCTTTCATATACATTAATGTT	1351
Db	673	GAAATGACCGCTTTAGGTAAAGCCTTAGGT-----GCAATGGGCAACTTTTATGGCC	726
QY	1352	CCATCATGCTCTTGAGATCTATATATACATGTACAAACAGAACATTCACGCAATATGGCT	1411
Db	727	ATGGCTGGTTTGGGTGATTTAGTTTAACTGATCTGACAAACCAATACAGTAAATGCTGCG	786
QY	1412	TTTGGACTAGAAATAGAAAAAGTATGAATATTAATATCATTAATATAGTCAACACTTAAG	1471

QY 2876 AATCCGGATATCTGTACATATTTCTACTAAGACGGTACTTAATGAGACATTT 2935
DB 668 AACTTAATATTTTCTGTTAAACACCTACTTAACAAATATTTATTTTACATACAAATTT 609
QY 2936 AATCTATATATA 2948
DB 608 TATTAATATCGTA 596

RESULT 24
US-10-204-708-59/c
; Sequence 59, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: FIEBENROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; PRIOR FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 59
; LENGTH: 6156
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-59

Query Match 2.0%; Score 94.6; DB 4; Length 6156;
Best Local Similarity 42.5%; Pred. No. 7.4e-09;
Matches 850; Conservative 0; Mismatches 1134; Indels 15; Gaps 6;

QY 351 ATAGAGTCTTATTCGACATACATATTAATCAATCAAAACAAATATATAGTATC 410
DB 2031 AATATATATCTTATCTTCTTCTTACAGCTCTTCAAAAAATTAATATCAATTAC 1972
QY 411 TAAATTAATTAATAATCCCTCAATAGCTTTAGCAAGACTCAAAATATATAGACACTT 470
DB 1971 TAAATTAATTAATAATCCCTCAATAGCTTTAGCAAGACTTTTAACTTACATTAACCTT 1912
QY 471 AACTTAATTAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 530
DB 1911 AAAACCATATCAACACTTCAATATATATATTAATTAATTAATTAATTAATTAATTA 1852
QY 531 CTGACATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 590
DB 1851 AAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1792
QY 591 TATAATTTTGAATGATTTCTGATATTAATTAATTAATTAATTAATTAATTAATTAAT 650
DB 1791 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1732
QY 651 GTGACGATTCATTTGGCAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTA 710
DB 1731 ACTCTTTCTAATCTTACCAAAACCTCACTTAATTAATTAATTAATTAATTAATTAAT 1672
QY 711 ACTTAAGGAGACGATGATATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 770
DB 1671 AAATACCAACAAATTTTCAACACCTTAACAAATTTTAAATTAATTAATTAATTAAT 1612

QY 771 AATATTTACCAATATATCTTACAGACATATATATGCAACGCAATATATAGCAAG 830
DB 1611 TATATATACCAACCCCTATATATATATATATATATATATATATATATATATATAT 1557
QY 831 TATATATACCAACAT 890
DB 1556 ATTTCTTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1497
QY 891 TATGACACAAATATACCAACAAACAGCATATGTAATTAATTAATTAATTAATTAAT 950
DB 1496 AATCCATCTATATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1437
QY 951 GTAAGGAT 1010
DB 1436 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1377
QY 1011 AATATATACCAATTTTATATCTCTGATCAAGTTTGTCAAGAAATTAATTAATTAAT 1070
DB 1376 AATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1323
QY 1071 TTCTTGTAT 1130
DB 1322 TACCTTATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1263
QY 1131 CAAT 1190
DB 1262 TAT 1203
QY 1191 GAGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1250
DB 1202 TACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1143
QY 1251 GAT 1310
DB 1142 TAT 1083
QY 1311 TACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1370
DB 1082 AATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1023
QY 1371 TAT 1430
DB 1022 TAT 964
QY 1431 AAGTAT 1490
DB 963 AAT 904
QY 1491 CTGTAAT 1550
DB 903 AT 844
QY 1551 CTAT 1610
DB 843 ATCTAT 784
QY 1611 CTAT 1670
DB 783 TAT 724
QY 1671 AAT 1730
DB 723 TAT 664
QY 1731 AGAT 1789
DB 663 TAT 604
QY 1790 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1849
DB 603 AT 544

QY 1850 AAATTAACCTTTACTATTATTAATAATATGTTGACACTTTTGCATGATCANG 1909
DB 543 ACCAAATTTTCTACTTATATATATCTTCAAAATTTTTCATCTCAAAATCTCAAAAT 484
QY 1910 TATTAATGAACATATTATTAATTAATAATTAATAAAGAACCAATATATATTA 1969
DB 483 AAATTAAT 424
QY 1970 -ATTACTAAACACCTTACTATAGTATATATATATATATATATATATATATAT 2028
DB 423 TCCCTTTATTTTCTATTTTCTATATATATATATATATATATATATATATAT 365
QY 2029 TATCATCTTGAATATTTTCAAAATATATATATATATATATATATATATATAT 2088
DB 364 TTTATATCTTAATATTTCCCAAAATTCATATACGATATATATATATATATATAT 305
QY 2089 TTCTTAATTAATTTGCTATTTACCCAAAGTTAAATATACATCTCCGCAATATG 2148
DB 304 CTCTAATCTTAAACATATATATATATATATATATATATATATATATATATAT 245
QY 2149 AAGAAATATACATATACATGTCATATATATATATATATATATATATATATAT 2208
DB 244 TACAAAT 185
QY 2209 AACATGAAACCAAT 2268
DB 184 TATCTTCAAAATATCAAAATCTCTATATATATATATATATATATATATATAT 125
QY 2269 AATTAAGAAATATGTTATGTTATATATATATATATATATATATATATATATAT 2328
DB 124 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 65
QY 2329 AACTCAAGTTCTTATTA 2347
DB 64 TTTCAAAACATATATATA 46

RESULT 25

US-09-543-681A-1742/C
Sequence 1742, Application US/09543681A
Patent No. 6605709
GENERAL INFORMATION:
APPLICANT: GARY BRETON
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
FILE REFERENCE: 2709.1002-001
CURRENT APPLICATION NUMBER: US/09/543.681A
CURRENT FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/128,706
PRIOR FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 8344
SEQ ID NO 1742
LENGTH: 1638
TYPE: DNA
ORGANISM: Proteus mirabilis
US-09-543-681A-1742

Query Match 2.0%; Score 93.8; DB 4; Length 1638;
Best Local Similarity 46.9%; Pred. No. 8.3e-09;
Matches 330; Conservative 0; Mismatches 367; Indels 6; Gaps 1;
QY 2727 TCATTAATTAATTCATTTTACATTTAGTATTTGTTAGTACCTGGCATTCAT 2786
DB 1357 TTAATAATACCGCGGCTCATATATTTTCAAGCTCTGTACATTTCCCGCAAGAT 1298
QY 2787 ATGATGATATGCTATTAAGTTTCACTTATATATATATATATATATATATATAT 2846
DB 1297 AATGGTATATGCTCCAAAGCATGTTGAATACAGTATATTTTTCATATGATCGGC 1238
QY 2847 TACAGATGCTATTTATTAATTTACATACATTTCCGATATATTTGACATATTTCTACTA 2906
DB 1237 GGGCAATTTTTCGTAATAATTTTGGCCATATAGGGAATATCTTCAAGTCTGCTCCGTA 1178

QY 2907 AAGAGGACTCTAATTGGAAGACATTTATCTATATATATATATATATATATATAT 2966
DB 1177 ATGGCGGAT 1118
QY 2967 CAGCTTTTACTTCACTTCAATATCTTTGGAAGAAGACATATATCTCATCTATAC 3026
DB 1117 CTGATCCGTCAGTTCATATATATATATATATATATATATATATATATATATAT 1058
QY 3027 TACAGGAATCTTCTATTTTCCCTATATATATTTTCCCTCTGTAATATGATATC 3086
DB 1057 GATATCTTTTATACACTTATCTTTCAATTTTACGCTCTTGGAAGAACCGTATATGTT 998
QY 3087 TTAATGCGATCGTATCTGTAATCTGTTATCTGATATATATATATATATATAT 3146
DB 997 TTGGCTGTAATTCATATGCGCATATCGCAATTTTCATGAGAAAAGCTATTTACTG 938
QY 3147 CTGCTCTATATTTCCATATATGAGAGTACTATATAGCAATATATATATATCTT 3206
DB 937 CCAATTCATATGACCCCATGAGCATGTTGACACCG-----TAAAGCGCTTTAT 884
QY 3207 CACTACCAAT 3266
DB 883 CATGTCAAAAGATATCACTTTCTAAGAACCGGAAGTACGTGCAATATCTATTTTA 824
QY 3267 TAAATGAGTATCATACCCCTTGATTTTATATATATATATATATATATATATAT 3326
DB 823 TCATGCTTTATTTTATAGATGACTTAAACGATGATCGCTGCAATATATCTTTAC 764
QY 3327 CAACACCTGGGAACAGATATATATATATATATATATATATATATATATATATAT 3386
DB 763 CAGTACCTGTTTCCGCTAAGTATATATATATATATATATATATATATATATATAT 704
QY 3387 TACTTCTAATTTACGTTAATACAGGAGTTACGACTATTTG 3429
DB 703 GTTCATGAT 661

RESULT 26

US-10-204-708-57/C
Sequence 57, Application US/10204708
Patent No. 6677731
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIERENROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
FILE REFERENCE: 5013.1012
CURRENT APPLICATION NUMBER: US/10/204.708
CURRENT FILING DATE: 2003-05-06
PRIOR APPLICATION NUMBER: PCT/EP01/03971
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: DE 10019058.8
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: DE 10019173.8
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 98
SEQ ID NO 57
LENGTH: 6326
TYPE: DNA
ORGANISM: Artificial Sequence
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-57

Query Match 1.9%; Score 92.2; DB 4; Length 6326;
Best Local Similarity 44.9%; Pred. No. 2.1e-08;
Matches 436; Conservative 0; Mismatches 528; Indels 7; Gaps 2;

QY 1423 AATAGAAAAGGTAGAAATATTAATACATTAATAGATCAACACCTTAAGCTTTGAAG 1482
DB 3104 AAAAAACCTTATATACACATTAATCAAAATCCCTCTCAAAATCCGAATACACATCCC 3045
QY 1483 AACGAGTACTGTAAACCACTGATATCATTAAGCAAAAAAATTAATGTAAGTACCAAT 1542
DB 3044 AACCAATCCACTCTTAATAATTAATTAATCAATTAATTAATTAATTAATTAATTAAT 2985
QY 1543 TTGCATATCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1602
DB 2984 TACTTAATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2925
QY 1603 CATATTAATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1662
DB 2924 CCAAAATCTACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2865
QY 1663 AATATAGAAATCAACGCTTCACTATTAATTAATTAATTAATTAATTAATTAATTAAT 1722
DB 2864 AAAAAAATCAACCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2805
QY 1723 TCCAACTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1782
DB 2804 AACCAAAATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2745
QY 1783 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1842
DB 2744 AAAAAATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2690
QY 1843 AATAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1902
DB 2689 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2630
QY 1903 CATCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1962
DB 2629 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2570
QY 1963 AATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2022
DB 2569 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2510
QY 2023 AATTCATCATCTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2082
DB 2509 CCAATTAATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2450
QY 2083 CATTAATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2140
DB 2445 AACCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2390
QY 2141 ACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2200
DB 2389 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2330
QY 2201 CATTAATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2260
DB 2329 AATTCATCATCTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2270
QY 2261 TATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2320
DB 2269 TAAACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2210
QY 2321 AACCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2380
DB 2209 ATCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2150
QY 2381 CATAGTATCA 2391
DB 2149 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2139

RESULT 27
US-10-204-708-31/C
; Sequence 31, Application US/10204708
; Patent No. 6677731

GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIRENEBROCK, Christian
APPLICANT: BERLIN, Kuit
TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
FILE REFERENCE: 5013.1012
CURRENT APPLICATION NUMBER: US/10/204, 708
PRIOR FILING DATE: 2003-05-06
PRIOR APPLICATION NUMBER: DE 10019058.8
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: DE 10019173.8
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 98
SEQ ID NO 31
LENGTH: 8093
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-31
Query Match 1.9%; Score 92; DB 4; Length 8093;
Best Local Similarity 45.7%; Pred. No. 2.3e-08;
Matches 811; Conservative 0; Mismatches 890; Indels 75; Gaps 11;
QY 744 AATTAACACTTAACGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 803
DB 2137 AATTAACACTTAACGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2078
QY 804 TATATGAACGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 863
DB 2077 AATCTTAACGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2018
QY 864 CTAATCTTAACGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 923
DB 2017 TATTAACACTTAACGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1958
QY 924 GTAATAATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 983
DB 1957 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1898
QY 984 CCAATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1039
DB 1897 CCAATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1838
QY 1040 CCAATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1099
DB 1837 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1778
QY 1100 AATTAACACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1159
DB 1777 AATTAACACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1718
QY 1160 TACATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1219
DB 1717 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1658
QY 1220 GCAATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1279
DB 1657 AATTAACACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1607
QY 1280 AATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1339
DB 1606 AATTAACACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1560
QY 1340 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1399

[illegible]

Oy	1967	TAAATTAAGTAAAACGACCTTTTACTAGTAAGTAATTATCAATAGATGTCCGACATAATTAAT	2026
Db	2079	TAAAAAGAAAAAACAATTACGAATCACCACCAACCTTCCTATATATAAAAABAAA	2020
Oy	2027	TCTATCATCTTGACATPATTTGCCAACATATAATATGATTAACCTGTTTTTTTAGCCCAATT	2086
Db	2019	ACAAATCTAAATTAATTAATTTATTAACCTTAACATCACATCTCAAAATTAATAAAATPAACCG	1960
Oy	2087	AGTCTTAATTAATAATGGTAATATPACCCAATAAGTTAAATACACATTCCTSCACAATA	2146
Db	1959	TATCTAATAAAACAATTATTCATA	1900
Oy	2147	TGAAGAAGATACATACACAGTCACAAATTAATAACAAATTTCTAATATATACACCTAC-T	2205
Db	1899	TAAAAAAATTTTATATCA--ACGTTCAAATTAATCTTATATGTCTACACATATACAT	1842
Oy	2206	CTAAACATGGAAACCAATAGTATATCATATCCATGAATATATCCCTACAGACATATCA	2265
Db	1841	ATATATATATATATTTATACACATATCATTTCCAAACATTAATATACCAATATTTATA	1782
Oy	2266	AATATATTAAGAAAAATCGTATATGATATTTATATACAAATTAACAAATTTCCACCTTACAA	2325
Db	1781	TATATAAACATCTAAAAAACAATAAATCTAATATATACCTAATATATAAAAAAAAAATA	1722
Oy	2326	CAAACCTCAAGTCTATATATTAACAAT--TGAAGATATATATATCTACCCACACAT	2383
Db	1721	CTTACATATAATATACATACAAATTTATATATTAATAATATCTAATAATATCTGCATCCAC	1662
Oy	2384	AGTATCAAAATTCATATATACATCATATCCTTTTAAAAAGATGACTTCA	2433
Db	1661	TACTATTAACCTTTATATCTTAAATTCCTGATCTAAAAAATATATATTTA	1612

RESULT_29
US-09-328-352-3585/c
; Sequence 3585, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 3585
; LENGTH: 1293
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-3585

Query Match	1.9%;	Score 90.4;	DB 4;	Length 1293;
Best Local Similarity	46.2%;	Pred. No. 3.4e-08;		
Matches 341;	Conservative 0;	Mismatches 391;	Indels 6;	Gaps 1;

Oy	2677	TACTATATCTACTCGTAAATCTTTTGGCTGATATCATCTTCCTTAGAGAGATTCATATATTA	2736
Db	1128	TAAATAGCTATGGGTAAGTCTTCGACATTCATCTTTCATCCATTTGCCATGGTTAA	1069
Oy	2737	AATCCATCTATACATTAAGTAATGTCTTAAGTTAATCCTGGCCATCATATGACTGCAT	2796
Db	1068	AGCATTAAGAGAGGTGTTTTTAAGTCACGAGACATTTCTCTGGCAGTCATATTCATTA	1009
Oy	2797	TGCTATTTAAAGCTTATCTCATATACATGATACAAACCTATTTTTTTTACAGATGCT	2856
Db	1008	AACCTGATCGCTTTTGAATGACATGTGAGGCAATGTGATGCAATGGGTGAGCTCTC	949
Oy	2857	ATTCATATAATACCTACACAATTCGGGTATCTCTACAAATTTCTACTAAAGACGTAC	2916
Db	948	TTTTTAATATGATGACAAATAAGAGAAATATGACATTTTCTCTCGAGAGGTGGAT	889
Oy	2917	TCTAATTTGGAAGACATTTATCTATATATATAATCTTCACAAAACCTACACGCTTTAC	2976

Db 888 GAAATTTGGAATTCAGCGTAGACGCTAGAGTAGATCATGACGGAAGCCAGCGCTAGC 829
Qy 2977 TTCACTTTCATATCTTTGGAAGAAGACACATATATCTCATCTATCTATCAAGGAAT 3036
Db 828 CTCCTCGGTAAGGTTATATGTTGCTGTATATATTCGACATATGGCTGATGCGTTT 769
Qy 3037 CTACTATTTTCCCTATATATTTTCCCTCCTGTATATATCTGATATCTTATTTGCT 3096
Db 768 TTCAACCAACCTGCGTAAATTCACCTGTTTCAACCCGTATATATTTAGCTTGAAG 709
Qy 3097 ATCGATGTAATCTGTACTTCTATCTATATATATAGTACCATGATTTGGTCTAT 3156
Db 708 CTCCTAAGAAATCTGCAATTTCTATCCAGAAACAGGCTTCGCTCGGACGTTAAA 649
Qy 3157 AATTCGAATATGAGAGGTAAGTCTATGAGACATATATTTACTTTCTCACTACCAA 3216
Db 648 CACGCTTTAT-----GGTCGCAATTTGCTCCGTAAGACCTTTCACTGCGCAAA 595
Qy 3217 TATATTAACGAATATATATAGCTGTAGATAGATAGATATATATATATATGAGT 3276
Db 594 GAGTTCATCTTTATATATATATGAAATTTAACTGCAAGTTATTCGCAAAAATCTT 535
Qy 3277 ATCATACCCCTGTGATTTTATGATATGAGCTAGCACTACTTCTTCCAAACCTGG 3336
Db 534 ATTTAGCTGTGCTATATATCATGATGATGCTTTTGCCACAAAGTCTTTACCGCTACGA 475
Qy 3337 CGAAGCAATATGATATAGACTCGATCGATGATGATGATGATGATGATGATGATG 3396
Db 474 TTCGCGCGGCAAGCACTGSAATCACTGATATGCACTTTGATTAATGAAACAT 415
Qy 3397 ATTACGTATTAACAGGGA 3414
Db 414 ACGATCATTTCTGTGA 397

RESULT 30
US-10-204-708-50/c
Sequence 50, Application US/10204708
Patent No. 6677731
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
FILE REFERENCE: 5013.1012
CURRENT APPLICATION NUMBER: US/10/204,708
CURRENT FILING DATE: 2003-05-06
PRIOR APPLICATION NUMBER: PCT/EP01/03971
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: DE 10019058.8
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: DE 10019173.8
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 98
SEQ ID NO 50
LENGTH: 6306
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-50

Query Match 1.9%; Score 89.8; DB 4; Length 6306;
Best Local Similarity 44.3%; Pred. No. 5,7e-08;
Matches 583; Conservative 0; Mismatches 722; Indels 10; Gaps 5;
Qy 1148 CTAATAATATATACCAATCAAGATTTATAGTGTACAGTGTGAGCTGATTAAGAAC 1207

Db 6212 CTATTAACCTCTTTCACATATATTAATAAGCTTTATCATTAATAAATCTTTATTTAT 6153
Qy 1208 ATAAATGCAATTCATGATGATATATATGCTGGAATAAATTTAGTATATATGCTTGGT 1267
Db 6152 ACACTTAACCCGATCTTATCTTAACTATTTCTATTCATAATATATCTTAAACCT 6093
Qy 1268 ACTGTTATATTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 1327
Db 6092 ATCTTAACCAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 6036
Qy 1328 ATAGATCTTCAATATATGCTGATGATGATGATGATGATGATGATGATGATGATG 1387
Db 6035 AACTATATATTTCAAT 5976
Qy 1388 ACGAATCTTCAAGCAATATGCTTTGACTGATGATGATGATGATGATGATGATGAT 1446
Db 5975 AATAACATTTCTATATATACCAAAACATCAATTAATAATCTTTAATTAATAAATA 5916
Qy 1447 TACATTAATGATCAACCTTAACTGTTGAAGAACGATGATGATGATGATGATGAT 1506
Db 5915 TGCATATATACCAAAACCAAAATATATATATATATATATATATATATATATATAT 5856
Qy 1507 ATCATTAAGCAAAATCTTAATGATGATGATGATGATGATGATGATGATGATGATG 1566
Db 5855 TATATTAATAAATAATTAATAAATCTTTAATAAATAAATAAATAAATAAATAAATA 5796
Qy 1567 ACATGAT 1626
Db 5795 ATATATTAATAAATTAATAAATCTTTAATAAATAAATAAATAAATAAATAAATA 5736
Qy 1627 TGTATATATTCACACATGATATATATATATATATATATATATATATATATATAT 1686
Db 5735 ATTTATATTAATAATTCACACATATTTTCTTAATAAATAAATAAATAAATAAATA 5676
Qy 1687 CATATATTAATAAGCATCAATATCCATGCTTATATATATATATATATATATAT 1746
Db 5675 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 5619
Qy 1747 CATTAACAACTTAAT 1806
Db 5618 AAAAAATTAATAAATCACTATATATATATATATATATATATATATATATATATAT 5559
Qy 1807 TCCATCTCAATTAACATATGTTATGATATATATATATATATATATATATATAT 1866
Db 5558 AATCCATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 5499
Qy 1867 ATTTATTAATAAT 1925
Db 5498 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 5439
Qy 1926 TTTTATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1985
Db 5438 TTAATAAATCTTATCAACATATATATATATATATATATATATATATATATATAT 5379
Qy 1986 TTTATAGTAAT 2045
Db 5378 TTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 5319
Qy 2046 TTCAACATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2105
Db 5318 CAAAAATATCAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 5259
Qy 2106 AATTACCAAAAGTTAAAT 2165
Db 5258 AATATTTAT 5199
Qy 2166 GTCAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2223
Db 5198 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 5139
Qy 2224 AGTATATATATCTTAAGTATATATATATATATATATATATATATATATATATATAT 2283

Db 5138 AATATAAACTCTTAAAAATCAATATAAAAAAATCAAAAAAATATAACAT 5079
Qy 2284 TTAATATATTTATACAAATTACAAAATTTCCAACTTCAACAAACTCAAGTTCTAT 2343
Db 5078 AATATACCAATTAATATCTAAAAAATATCTCAAACTTTCAATATCAAAAAATCAAAAT 5019
Qy 2344 TATAAATATCAAGATATATATATATCTTCAACCAACATATGATATCAATCTATATA 2403
Db 5018 TAAAAATCAGACAAAAAACAATCTTATCTCAATCAATTAATAAAAAAATCTCCAAAAT 4959
Qy 2404 CATCATCTCTTTTAAAAAGATGACTTCACAATTCATATCAATTAATATATACA 2458
Db 4958 TACCAAAATATATACATATTAATCTTATATCTTATATCTCAAAATTAATAATAA 4904

RESULT 31
US-10-204-708-10/c
Sequence 10, Application US/10204708
Patent No. 6677731
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIRENBROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
FILE REFERENCE: 5013.1012
CURRENT APPLICATION NUMBER: US/10/204,708
CURRENT FILING DATE: 2003-05-06
PRIOR APPLICATION NUMBER: PCT/EP01/03971
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: DE 10019058.8
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: DE 10019173.8
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 98
SEQ ID NO 10
LENGTH: 6070
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-10

Query Match 1.9%; Score 89.6; DB 4; Length 6070;
Best Local Similarity 44.2%; Pred. No. 6.2e-08;
Matches 612; Conservative 0; Mismatches 749; Indels 23; Gaps 5;

Qy 1130 ACAATATAGATGATGTTCTTAAATATATATATCCATCAAGATATATAGGTACAGTT 1189
Db 2957 ACTAATATAAATCGCTCTCAAAAAAATTAATTAATTAATTAATTAATTAATTAATTAAT 2898
Qy 1190 GGAGCTGATTAAGACATATATGCAATGATGATGATGATGATGATGATGATGATGATGAT 1249
Db 2897 AAAAT 2838
Qy 1250 GGTATATATGCTGTTCTATCTATATATATATATATATATATATATATATATATATATAT 1309
Db 2837 TAT 2778
Qy 1310 ATAGCAAAAT 1369
Db 2777 AACAT 2718
Qy 1370 CTAT 1429
Db 2717 TAT 2658
Qy 1430 AAGGTAGAAAT 1489

Db 2657 AT 2604
Qy 1490 ACTGTAAAAACATGAT 1549
Db 2603 TCCCTTAT 2544
Qy 1550 TCAAT 1609
Db 2543 TTTTAAAAAT 2484
Qy 1610 TCTTAT 1669
Db 2483 TCAAT 2424
Qy 1670 GATATACAGCTTCAACAT 1729
Db 2423 TAAAAATCCCATTTCTACATCTTTACACTAAAAAATATATATATATATATATATATATAT 2364
Qy 1730 -----AAGTAT 1781
Db 2363 ATCTAT 2304
Qy 1782 CTAT 1841
Db 2303 AAAAAAAT 2244
Qy 1842 AAT 1901
Db 2243 AAAAA--TAT 2187
Qy 1902 ACTAT 1960
Db 2186 ATAAAAAT 2127
Qy 1961 AT 2020
Db 2126 AT 2067
Qy 2021 AT 2080
Db 2066 AT 2007
Qy 2081 CCAT 2140
Db 2006 TAT 1947
Qy 2141 ACAAT 2195
Db 1946 AAAAAAT 1887
Qy 2196 ACCACCTAT 2255
Db 1886 AAT 1827
Qy 2256 GCAT 2315
Db 1826 CTAT 1767
Qy 2316 ACTTCAACAAACAAATCTCAAGTATATATATATATATATATATATATATATATATATATAT 2375
Db 1766 AC 1707
Qy 2376 CAACAT 2435
Db 1706 AAAAAAAT 1647
Qy 2436 ATTCAT 2495
Db 1646 TAT 1587
Qy 2496 TTTT 2499
Db 1586 ATAT 1583


```

QY 1790 ATAAACAAATTCCTATCTCTATCTCA-ATTACCAATGTATACAGATTAAATAGT 1848
DB 2048 AAAAAAAAAATTATATAAATACTAATCCGACAAATCCCTAATTATTAACCTTTAAATTCAA 1889
QY 1849 CAATTAACCTTTACTATTTATTAATAATATATGTGACACCTTTTTCATCATCAT 1908
DB 1988 AATATAATACATTTTATTTAAATAAATAATCTTCAAAACAAATATCTAAATATC 1929
QY 1909 GTATTAATGAACATTAATTTTACATTAATAATACATTAATAAAGAACCAATCATATTTA 1968
DB 1928 ACCATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1869
QY 1969 AATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2028
DB 1868 AATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1809
QY 2029 TATCATCTTGACATTTTCAAACTAAATATGATTAAGTGTGTTTATGCTCCATTAG 2088
DB 1808 TCAAAATAAT--AAATTAACAACATTAATAATTTTAAAC--TTTATCTCTCAAAATAT 1754
QY 2089 TCTTAATTAATTTGCTATTTACCAAAAGGTTAAATACACATTCCTGCAACATATG 2148
DB 1753 AATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1686
QY 2149 AAGAAGATACATACACGTCAAATTAATTAACAAATTTCTAATATTAACAACCTACTCTA 2208
DB 1695 -AAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1637
QY 2209 AACATAGAACCAATAGTATATCATCTAGTAAAGATATCCCTACAGCATATTCAAAT 2268
DB 1636 AATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1577
QY 2269 AATTAAGAAAAATCGTATGATATTTATTAACATTTACAAATTTCCAACTTCAACAA 2328
DB 1576 TAACCTAAATATACACTAATATTAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1517
QY 2329 AACTCAAGTCTATTAATTAACAATTCAGAAATTTATATCTTCACCAACATATGTA 2388
DB 1516 AAAAACTAATATATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1457
QY 2389 TCAAAATTCATATATCATCATATCTTTTAAATAAAGATGATCCATTCATCATATAA 2448
DB 1456 TAAAAATATACGTCCCAAAATTTTAAATTAATTAATCACTAATAAATAAATAAATAAATAA 1397
QY 2449 TAATTAATACATCAATGTTTATTCGA 2475
DB 1396 TTTCTCTCTTATACATTTTTCAAA 1370

```

RESULT 33
US-10-204-708-12/c

```

; Sequence 12, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; PRIOR FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019056.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98

```

```

; SEQ ID NO 12
; LENGTH: 6317
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-12

Query Match 1.9%; Score 88.2; DB 4; Length 6317;
Best Local Similarity 43.5%; Pred. No. 1,1e-07;
Matches 925; Conservative 0; Mismatches 1163; Indels 39; Gaps 10;

QY 2295 TCAAGAATATACCTCTCTCTGTCATCCATATTTTCCATTTATTTACATATCATG 354
DB 3714 TTCAAAATTCATCCATATATATATATCAATTAATTTCTTTATTTATTAATAAATAATAC 3655
QY 355 AGTCTTTATTCGACATACATATTAATCAATCAAAACAAATATAGTACTATCTAT 414
DB 3654 TACATTAATATCATTTATTTATTCACATCAACAAACAAATACATATTAACATTA 3595
QY 415 TATATAAATAATCCATATAGCTTTAGCAAGAACATCAATATATAGACATTA--A 472
DB 3594 ATAAAAATTAATACAACTATATTTTAATATTAATATTAATTAATTAATTAATTA 3535
QY 473 CTATTAATATATTTCAAAATTTATTAATTAATTAATTAATTAATTAATTAATTA 532
DB 3534 ACATTAATTAATAAACAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTA 3475
QY 533 GACATATTAACAATAGTATACATTTCTTAACATGATTAATATACAAATATATATA 592
DB 3474 AAAAACTAAACAAATATTTCTTAATTAATTTCAAAATTTCAAAATTTTAATTA 3415
QY 593 TAATTTTACATGTTTCTGATTAATTAATTAATTAATTAATTAATTAATTAATTA 652
DB 3414 TACCTAAACATTAATAACAAATAATTTTCTTCTCTCTTCAACCAATTTCTAA 3355
QY 653 GCAGATCATTTGGCAGACCAATAGCAATTTGACATGACGACAT----GGTATATCAG 707
DB 3354 ATACTTAATATTCCTCATACATATTAATTAATTAATTAATTAATTAATTAATTA 3295
QY 708 TTAATTAATGGGACGTGATCATAGAAATTAATTAATTAATTAATTAATTAATTA 767
DB 3294 TATATAATACACATACACACACACCCCTTAATTAATTAATTAATTAATTAATTA 3235
QY 768 TAAAAATTTACCAATATCATCTACGACGAACATATATGCAACCCGCAATATAGAG 827
DB 3234 TCAAAATTAATCTTAACAT-----TCAATAAATAATTAATTTTATTAATTA 3181
QY 828 AAGTATTAATCGCAACAATATACATATATATCTTAATCTTACACAAATTAACGA 887
DB 3180 AATTAATTAACGACAACTTAATTTTAAATAATATATATACCTCCCAATTAATTA 3121
QY 888 CCAATATGACAAATATACAAACAAACGACATATGTAATAATCTCCAAATTAATTT 947
DB 3120 TCTATTTTAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3061
QY 948 GTAGTAAGGATTCGAATTAATTAATCACTCAAAATTTCCAGTAATAGCAGAGAAATTT 1007
DB 3060 AATTAATAAGTTTCTTAATAAATAAATAATTAATTAATTAATTAATTAATTAATTA 3001
QY 1008 TACATTAATTAATCAATTTTAAATCTCTGTCGCAAGTTTTCCTAAAGAAATTTGCA 1067
DB 3000 TCTATTTTCTTAAACAAATTAATCTTACAGCCCTCTTAATCTTATCTTACAAAC 2941
QY 1068 ATCTTCTTGATATAGTATGCTGTCGATATTAAGAACTTGGTGAATCATTTATAG 1127
DB 2940 ACAAAC-----ATAATTCCAAATTTCTCAAAATAAATAAATAAATAAATAA 2885
QY 1128 AATCAATTAATGATATGTTCTTAATAATATATATACATCAATATTAATTAATTA 1187
DB 2884 AATCAAAATAATATATTTTACATATTAATAAATTTACATATTAATTAATTAATTA 2825
QY 1188 TTGAGCTGATTAAGAATATTAATGCAATTTGCAATGTAATTAATCGCTGAAAAATTT 1247

```


Db	2824	TTCTTACCCAAAAACAAAAATAATATATTAATAAAAGTCAATATATAAATAATTAATTCCT	2765
Qy	1248	TAGGTAAATTAATGCTGTGTCTACTGTATTACTA-AAGCGATGANTGAATTAATAACCTA	1306
Db	2764	AAAATAATTAACCCATCCTACATTAAATAATAATATACTTTAATAATAAACACTTAATCTC	2705
Qy	1307	TATATAGCAAAAAATCATTCATATGATCTTCATACATTAATTGGTCATCTGTCTTGA	1366
Db	2704	TCATATTTCAATACCCCTTATTTATAAACAATAATATATTTCTCTTCATATACATA	2645
Qy	1367	GATCTAATATTAAATGACAGCAAGAACATTCAGCGAATAT-----GGCTTT	1413
Db	2644	CAAAAACATTTCAAACTAAACATAAAAATATGCAATCAATTAATATTAATTTATCATA	2585
Qy	1414	TGACATAAATAGGAAGAAGGTAGAAAAATATAA-TACATTAATGATCACACCTAAAGC	1472
Db	2584	TCATATCCAAATATTAACAAACCAAAATTTATTAATTAATTAATTAACAAACCTCCAAAC	2525
Qy	1473	TTGTTGAAGAACCAATCTGTAAACCACTGATATCTTTGCAAAAAACTTAATGTAG	1532
Db	2524	ATTAATAAAAAACTCCCAATTAATATCCAAATACTTAATTAATTTCTACTATATTC	2465
Qy	1593	AACATACCAATTTGCATATCTATTTACATTTATTAATACATGAAATATACCTAGATAAG	1592
Db	2464	CAACATTTATCTTAATTTCTTAACAAAATTTATTCATTAACATCTCCAAACACTTTAA	2405
Qy	1593	CCATATCAACATATTATATCTATGTCATCTTGTATTTTCAATTTCAACACATGATTAAT	1652
Db	2404	TAAATTTCTCATTTTACAAACAAAAAATTAATAAAACAAATAACTTAATTTCCCATATA	2345
Qy	1653	AAAGGTCAAAAATTTAGATATCAACAGTCCACCAATATATAAAGAGTCAAAATACC	1712
Db	2344	ACAAAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTCCT	2285
Qy	1713	ATCGTTAACATCCAACTAAGTATGTAACTATTACATAACAACTTAAGTAACTTAA	1772
Db	2284	TTTTTTCTATCAAAAAAATTAATAAAAAAATAAATAAATTAACCTTCATTAATTAACAACATTT	2225
Qy	1773	ATTTCCGTAGTAAATTAATTAATAAACAATTCCTATCTCGATGCAATTTAA-CCAAATGTTA	1831
Db	2224	ATTAAACACATCAATATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTTTC	2165
Qy	1832	TCAGATTATTAATAAGTCAAATTTAACTTTACTTATTTATTAATAATAATATGTTGACA	1891
Db	2164	AATTAATTAACATATTTCTTAATTAATAATTTCTTTAATTTTAATTAATTAATTAATAATA	2105
Qy	1892	CTTTTCACTACATCATGTATTAATTAAGAACATTAATTTACATTAATAACTAATAAAG	1951
Db	2104	AAATTAACACTTCATAAAAATCATACCTATCAAAATTAATAAAATCTTCAAAAGCTCTAATTT	2045
Qy	1952	AACCAATACATATATTAATTAATTAATTAACCTATTTCTAGCTAATAATTACATGATGT	2011
Db	2044	AAACTTCATTTCTATCACTATTAATAAGCCCAAAATTAATTAACGCAATTTTACCCGCATA	1985
Qy	2012	CCTGACATTAATTAATTTCTATCATCTTGACATATTTCAAACATTAATAATGATTAACGT	2071
Db	1984	ACATTAACCTTAATTAATTAATTAACAAATTTCAAAATTTCTATTAATAATTTCTTTCAATACAT	1922
Qy	2072	TTTTTTATAGCCATTAAGTTCCTTAATTAATAATGCTATTAATTAACCAAAAGGTTAAATAAC	2131
Db	1924	CATACCTTAATTAATTAATTAATTAATTAAGTAATCATTTATTAACAAATCATATTAATAA-----	1877
Qy	2132	ATTCTGCAACAATTAAGAAAGATACATTAACGTCAAATTAATTAATAACAAATTTCTAA	2191
Db	1869	ATATGCCCACTACTTAATAAAAAAATCAAAACCAATCACTACCAATTAATTAATTAATTA	1810
Qy	2182	TATAACCACTACTTAACATAGAACCAATAGATATCATACCTAACTAAGTAAATATCC	2251
Db	1809	TTCAATTAATTTTCAAAAAATTAATTAATAATAAGTA-AAACGTAACTAAGATCTACAT	1751
Qy	2252	TACAGCATGTACAAATTAATAAGAAAAATCGTTATGTATATTTATCAATTCAAATAAT	2311

Db	1750	AAAAAATCCTTCACAAAAAAATTCATTAATAAAAAAATCGCTACAAATTCCTAC	1691
Qy	2312	TCGAATTCACACACAAACATCAGTCTATATTAACAATTCAGAAATATATATCT	2371
Db	1690	TCATCTCTCCAAATATATCTATATATTCCTTAATATAACATTAATAAATTTTACATTA	1631
Qy	2372	TCACCAACACATAGTTATTCAAATTCGA	2398
Db	1630	ACACAAAAAAGATTCACAAATTCGTA	1604
RESULT 34			
US-10-204-708-36/c			
: Sequence 36, Application US/10204708			
: Patent No. 6677731			
: GENERAL INFORMATION:			
: APPLICANT: OLEK, Alexander			
: APPLICANT: BERLIN, Kurt			
: TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication			
: FILE REFERENCE: 5013.1012			
: CURRENT APPLICATION NUMBER: US/10/204,708			
: CURRENT FILING DATE: 2003-05-06			
: PRIOR APPLICATION NUMBER: PCT/EP01/03971			
: PRIOR FILING DATE: 2001-04-06			
: PRIOR APPLICATION NUMBER: DE 10019058.8			
: PRIOR FILING DATE: 2000-04-06			
: PRIOR APPLICATION NUMBER: DE 10019173.8			
: PRIOR FILING DATE: 2000-04-07			
: PRIOR APPLICATION NUMBER: DE 10032829.7			
: PRIOR FILING DATE: 2000-06-30			
: PRIOR APPLICATION NUMBER: DE 10043826.1			
: PRIOR FILING DATE: 2000-09-01			
: NUMBER OF SEQ ID NOS: 98			
: SEQ ID NO 36			
: LENGTH: 9347			
: TYPE: DNA			
: ORGANISM: Artificial Sequence			
: FEATURE:			
: OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)			
US-10-204-708-36			
Query Match 1.9%; Score 88.2; DB 4; Length 9347;			
Best Local Similarity 42.6%; Pred. No. 1.2e-07;			
Matches 918; Conservative 0; Mismatches 1213; Indels 22; Gaps 8;			
Qy	336	ATTATTTACATATCATAGAGGTTTATTATTCGACATCATATTAATTCATCAAAACAA	395
Db	8754	AATCGAAAAAATTCATTCATCTATAAACAACAAAAATATATATCATTAATTAATTA	8695
Qy	396	AATATAGTACTATCTATTAATTAATTAATTAATTCCTCAATAGCTTTAGCAGAACTACAA	455
Db	8694	CAAAAACAAATCTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	8635
Qy	456	TATATAGGACACTTAATCTATTAATTAATTTTCAAAATTTTATATAGTACATCATATTT	515
Db	8634	CTTAATAAATAAATAAATAAATAATATACAAAATCTATTTAAATACCTAAATAATTCATTTCT	8575
Qy	516	AGCATTAATTCACATCTGCACATGAATTAACAATAGTAACACATTTCTTAACATAGATAT	575
Db	8574	CCATAATATATATAAATAATTTAAATCTCTATCAAAATTAATAAATAAATAAATAATATCT	8515
Qy	576	ATACAAATTAATTAATTAATTTTGAATGTTCTTGATTAATTAATTAATTAATTAATTAATTA	635
Db	8514	ATAATAAATAATCTTAATAAATTTTAAATATTTCTTAATAAATAAATAAATAAATAAATAA	8455
Qy	636	AAATCAGTATTTTAGTGACAGATCATTTGGCAGACAATAGCAATTCAGCTGTACGAC	695
Db	8454	TCAACACATATATCCCAACACTTTAAATAAACAACAAAAAATTTCTTAAACCAAA	8395
Qy	696	ATGATATATCGTTAATCTATATGGGACGTGATCATATGAATATATATACATTAACACTT	755

Db 8394 ATTCCAAACAAACCTAAACAAACAAACCTTATCTATTAATAAAAAAATAAT 8335
Qy 756 ACCGAAAAATTTAAATATTTACCCAGATATCATCCAGAGACAAATATGCAACCA 815
Db 8334 AACCAAAATATGATATATATATATCCCAACCTTAAAAAAACGAAACATATTAACAC 8275
Qy 816 GCAATATGAGCAAGATATATATCTGCAACAAATGATATATCTTAATCTCTACAC 875
Db 8274 CTAATATCAAAAACCTTAAACCAACCTTACCAACATATTAACCCATCCCTACTAAA 8215
Qy 876 AACAAATGACCAATATGTAACAATAACAACAACAGCATAT---GTGTAAT 931
Db 8214 ATACAAATCTACTAAATATATATATTAACCTTACATCCCACTCTGAAAAACTTA 8155
Qy 932 ACTCCATATATATATGTAAGGTATGAAATTCATCATCAATTTCCCACTGAA 991
Db 8154 AACATTAATATGCTTCAATCTTAAACGAAATTCATTAACCGAATCCGCTCCACT 8095
Qy 992 ATAGCAGAAATTTTAAATATATCAATTTTATCTCTGCTCCAGTTTGTCT 1051
Db 8094 ACATCCAACTTAAACAAACAAATTAACATATCTCAAAAACAAACAA---AACA 8039
Qy 1052 AAAGAAATGAGCAATCTTCTGCTGATAGTATGCTGCTGATATTAAGAACTT 1111
Db 8038 AAAACAAACAAACAAACAAACAAATATATATATGCTACTTAAATCCCG---CTA 7983
Qy 1112 GGTGATCATGATATGAAACATATATATATGATGCTTAAATATATATATATATAT 1171
Db 7982 CTAAATAATCTAAAT 7923
Qy 1172 ATTAAGGTATGATGAGTCTGCTGATTAAGAAATATATATATATATATATATAT 1231
Db 7922 AATCAACCACTTACCTTCAACCTTAAACAAACAAATATATATATATATATATAT 7863
Qy 1232 ATCGTGAACAAAT 1291
Db 7862 AATATATATATCTTCAACAAATTTTAAATAATATATATATATATATATATATAT 7803
Qy 1292 GAAATTAACAACTAT 1351
Db 7802 ACTAAATATCTTAAAT 7743
Qy 1352 CCATGATGCTGAGATCTATATATATATATATATATATATATATATATATATAT 1411
Db 7742 TPAACATATATATCTCTCAATATATATATATATATATATATATATATATATAT 7683
Qy 1412 TTGGAATGAAATATGAAATATGAAATATATATATATATATATATATATATAT 1471
Db 7682 CCACCTCTTAAATAATATATATATATATATATATATATATATATATATATAT 7623
Qy 1472 CTGTGGAAGCAACAGTACTTAAACAACTGATATCATTAAGCAAAAATTAATGTA 1531
Db 7622 ACTTCTTAAACAACTCTCTCTTCAAAAATTAATATATATATATATATATATAT 7563
Qy 1532 GAACATCAATTTGAT 1591
Db 7562 TATTTCTTTTCCCTTTTCAATTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 7503
Qy 1592 GGCATATCAACAT 1651
Db 7502 TTTTAAATTTTAT 7443
Qy 1652 TAAAGGCTGAAAAAT 1711
Db 7442 TCTAATATCAAT 7385
Qy 1712 CATGCTTACATCAACATATATATATATATATATATATATATATATATATATAT 1771
Db 7384 ATTTAATGCAAAACAT 7335
Qy 1772 AATTTCTAGCTAT 1831
Db 7324 TAACCTTAATCACTATCTCT---AAATATATATATATATATATATATATATAT 7267

Qy 1832 TCAGATTAATAAATAGTCAAAATTAACCTTTACTATTTATTAATAATATATGTTGACA 1891
Db 7266 TTTTATCTCAACTACTATATATTTTAAACCTTATACATATCAAAATCAAAACCACTTA 7207
Qy 1892 CTTTTCATCATCATGAT 1951
Db 7206 ATTACTCTCTCAACAAATATATATATATATATATATATATATATATATATATAT 7150
Qy 1952 AACCAATCAT 2011
Db 7149 AATTAACAAATAAT 7090
Qy 2012 CCGACAT 2071
Db 7089 TCACAT 7030
Qy 2072 TTTTATGCCAT 2131
Db 7029 CTATTTATCAAAAT 6971
Qy 2132 ATTCCTGCAACAT 2189
Db 6970 CTAAAT 6911
Qy 2190 AATATACCACTACTCTTAAACATATATATATATATATATATATATATATATAT 2245
Db 6910 AAAAT 6851
Qy 2250 CCGACAT 2309
Db 6850 CCATATCAAAATAAT 6791
Qy 2310 TTCCACTTCAACAAACAACTTAAATGATATATATATATATATATATATATAT 2369
Db 6790 AAAACAAAT 6731
Qy 2370 CTTACCAACAT 2429
Db 6730 ATTAATAATCAACAT 6671
Qy 2430 TTCACATTCAT 2482
Db 6670 TATTAACCCCAAAATAATATATATATATATATATATATATATATATATATAT 6618

RESULT 35
US-10-204-708-35/c
; Sequence 35, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; PRIOR FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 35
; LENGTH: 9347
; TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-35

Query Match 1.8%; Score 86.8; DB 4; Length 9347;
Best Local Similarity 43.1%; Pred. No. 2.2e-07;
Matches 597; Conservative 0; Mismatches 767; Indels 20; Gaps 3;

1150 AAAATATATATACATCAAGATATATATAGTACAGATGAGCTGCAATTAAGAACAT 1209
DB AAAAT 9276
1210 AATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1269
DB AATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 9216
1270 TGTAT 1329
DB AAT 9156
1330 AGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1389
DB AGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 9096
1390 AGAATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1449
DB AGAATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 9096
1450 AT 1509
DB AT 9035
1510 ATTGCAAT 1569
DB ATTGCAAT 8975
1570 TGAAT 1629
DB TGAAT 8915
1630 AT 1689
DB AT 8855
1690 AT 1749
DB AT 8797
1750 AAT 1809
DB AAT 8737
1810 TATCTCAAT 1869
DB TATCTCAAT 8677
1870 TAT 1929
DB TAT 8617
1930 ACAT 1978
DB ACAT 8557
1979 ACAT 2038
DB ACAT 8497
2039 GACAT 2098
DB GACAT 8437

2099 AATGCTAT 2158
DB AATGCTAT 8377
2159 AATGCTAT 2218
DB AATGCTAT 8317
2219 CCAT 2278
DB CCAT 8257
2279 AATGCTAT 2338
DB AATGCTAT 8197
2339 TCTAT 2391
DB TCTAT 8137
2392 AATGCTAT 2451
DB AATGCTAT 8077
2452 TAT 2511
DB TAT 8017
2512 TAT 2571
DB TAT 7957

RESULT 36

US-10-204-708-22/c
Sequence 22, Application US/10204708
Patent No. 6677731

GENERAL INFORMATION:

APPLICANT: OLSK, Alexander
APPLICANT: PIEBENROCK, Christian

TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication

FILE REFERENCE: 5013.1012

CURRENT APPLICATION NUMBER: US/10/204,708

PRIOR FILING DATE: 2003-05-06

PRIOR APPLICATION NUMBER: PCT/EP01/03971

PRIOR FILING DATE: 2001-04-06

PRIOR APPLICATION NUMBER: DE 10019058.8

PRIOR FILING DATE: 2000-04-06

PRIOR APPLICATION NUMBER: DE 10019173.8

PRIOR FILING DATE: 2000-04-07

PRIOR APPLICATION NUMBER: DE 10032529.7

PRIOR FILING DATE: 2000-06-30

PRIOR APPLICATION NUMBER: DE 10043826.1

PRIOR FILING DATE: 2000-09-01

NUMBER OF SEQ ID NOS: 98

SEQ ID NO 22

LENGTH: 11049

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-22

Query Match 1.8%; Score 86.2; DB 4; Length 11049;
Best Local Similarity 45.9%; Pred. No. 2.9e-07;
Matches 513; Conservative 0; Mismatches 513; Indels 11; Gaps 4;

1530 TGAAT 1589
DB TGAAT 10868

```

QY 1590 AAGCATATCAACATATATCTTGTGTCATCCTTGTGTAATTTCAACACATATA 1645
DB 10808 TAAATATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 10749
QY 1650 AATTAACGGTCAAAATATAGATATCAAGTTCACCAATATATTAAGAGCAATATA 1709
DB 10748 CCTTACAAAAAATATATATATATATATATCA-TTTTCTTAAAAAATATTTTCTT 10690
QY 1710 CCCATCGTTAATCACTCAAGTATGATGATATTAATTAACAACTTAAGTAACTT 1769
DB 10689 AAAAAACAAAAACAAAAAATTTATTAATTTATTTACTTCACTTAA--- 10633
QY 1770 AAAATTTCTAGCTATATATATTAACAAATTCATTCCTATCTCAATTAACCAATGT 1829
DB 10632 AACTTTAAAAAATTTCTATCAATTAATTTATTTTCTTCTTTTATTTATTAATAT 10573
QY 1830 TATCAGATTATTAATTAAGTCAATTTAACTTTTACTTATTTATTAATATGTTTGA 1889
DB 10572 TTTTAACTTTTATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 10513
QY 1890 CACTTTTCACTACATCATATATATATATATATATATATATATATATATATATAT 1949
DB 10512 CAACTTCTCACTAATATATATATATATATATATATATATATATATATATATAT 10453
QY 1950 AGAACCAATATATATATATATATATATATATATATATATATATATATATATAT 2009
DB 10452 -----CAATATATATATATATATATATATATATATATATATATATATATAT 10398
QY 2010 GTTCGCACTATATATATATATATATATATATATATATATATATATATATATAT 2069
DB 10397 AATTTATTTCTTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 10338
QY 2070 GTTTTATATGCTTATATATATATATATATATATATATATATATATATATATAT 2129
DB 10337 ATTTAATTTATATATATATATATATATATATATATATATATATATATATATAT 10278
QY 2130 AATTTCTGCAATATATATATATATATATATATATATATATATATATATATATAT 2189
DB 10277 AAAACCAAAAAAATATATATATATATATATATATATATATATATATATATATAT 10218
QY 2190 AATATACCACTATATATATATATATATATATATATATATATATATATATATAT 2249
DB 10217 AACTTAAAAAATATATATATATATATATATATATATATATATATATATATATAT 10160
QY 2250 CCTACGATGATATATATATATATATATATATATATATATATATATATATATAT 2309
DB 10159 TCTTATATCTATATATATATATATATATATATATATATATATATATATATAT 10100
QY 2310 TTTCAACTTCAACAAAACTCAAGTCTATATATATATATATATATATATATATAT 2369
DB 10099 AATTTTAAATTTATATATATATATATATATATATATATATATATATATATATAT 10040
QY 2370 GTTCAACCAATATATATATATATATATATATATATATATATATATATATATAT 2429
DB 10039 CTTTCTTTAAAAAATATATATATATATATATATATATATATATATATATATATAT 9980
QY 2430 TTCACAATTCATATATATATATATATATATATATATATATATATATATATATAT 2489
DB 9979 ATCTTAAATATATATATATATATATATATATATATATATATATATATATATAT 9920
QY 2490 CAAGAAATTT 2498
DB 9919 TATATATAT 9911

```

RESULT 37
US-10-204-708-26/c

; Sequence 26, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt

```

; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; PRIOR FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019056.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 26
; LENGTH: 6583
; TYPE: DNA
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-26

Query Match 1.8%; Score 86; DB 4; Length 6583;
Best Local Similarity 45.0%; Pred. No. 2.9e-07;
Matches 490; Conservative 0; Mismatches 585; Indels 13; Gaps 4;

QY 1417 ACTGAATATAGGAAAGGTAGAAATATATATATATATATATATATATATATATAT 1476
DB 6526 ATTTAATATCCAAATATATATATATATATATATATATATATATATATATATATAT 6467
QY 1477 TGAAGAACAGTACTGTATTAACCACTGATATATATATATATATATATATATATAT 1536
DB 6466 CTCACATTAATTAATATATATATATATATATATATATATATATATATATATATAT 6407
QY 1537 ACCAATTTGATATATATATATATATATATATATATATATATATATATATATATAT 1596
DB 6406 ATCTATTAATTAATATATATATATATATATATATATATATATATATATATATATAT 6347
QY 1597 ATCAACATATATATATATATATATATATATATATATATATATATATATATATAT 1656
DB 6346 ATATACCTTAATATATATATATATATATATATATATATATATATATATATATATAT 6287
QY 1657 GGTCAAAATATATATATATATATATATATATATATATATATATATATATATATAT 1716
DB 6286 AATCAAAAAAATATATATATATATATATATATATATATATATATATATATATATAT 6227
QY 1717 TTAACATCAACTATATATATATATATATATATATATATATATATATATATATATAT 1776
DB 6226 ACATATATATATATATATATATATATATATATATATATATATATATATATATATAT 6167
QY 1777 CTAAGCTATATATATATATATATATATATATATATATATATATATATATATATAT 1836
DB 6166 ATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 6115
QY 1837 TTAATATATATATATATATATATATATATATATATATATATATATATATATATAT 1896
DB 6114 CTATATATATATATATATATATATATATATATATATATATATATATATATATATAT 6055
QY 1897 TCATCATCATATATATATATATATATATATATATATATATATATATATATATATAT 1956
DB 6054 CAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 5996
QY 1957 ATACATATATATATATATATATATATATATATATATATATATATATATATATATAT 2016
DB 5995 ACATTTATATATATATATATATATATATATATATATATATATATATATATATATAT 5936
QY 2017 CATATATATATATATATATATATATATATATATATATATATATATATATATATAT 2076
DB 5935 TAGGAATATATATATATATATATATATATATATATATATATATATATATATATATAT 5879
QY 2077 TATGCCATATATATATATATATATATATATATATATATATATATATATATATATAT 2136

```

Db	5878	ATTACTTAATAATTACTTTCTATATATATATTTTACATATATAATATAATATA	5819
Qy	2137	TGCACAATATGAGAGATACATACACGTCACATTAATTAACAAATTTCTAATATPA	2196
Db	5818	TAATACATATTTACTATTTATATATATATATATATATATATATATATATATAT	5758
Qy	2197	CCACCTACTCTAAACATNGAACCAATGATATATCATACCTAAGTAAGTATCCCTACAG	2255
Db	5758	TTTTTTTTTCAAAAAATTTTACGATACACAAAAACAAAAATTAAAAAAATACAAAAA	5699
Qy	2257	CATGATCAAAATTAATTAAGAAAAATCGTTATGATATTTATACATTTCAAAATTTCCA	2318
Db	5688	TTATTTTAAAAATTAATAAAAAATTAATCATATTAAGAAAAACTCACTTCTCATATTTACGATT	5633
Qy	2317	CTTCAACACAAACCTCAAGTCTATTTATTAACAAATTCAGAAATTATATACTTCAAC	2376
Db	5638	ACTACAAAACACCTATCTATATTTACTTATATGATTTTCACTTAATTTACTTAA	5579
Qy	2377	AACACATGTTATCAAATTTCTATATATACATCATATCTTTTAAAAAAGTAGCTTCA	2433
Db	5578	AACCAATTTCAATTATTAACATATATACCTTACCATTAATAAAAAATACCGTAAAAATAC	5519
Qy	2437	TTCATACATAATAATTTATACATGCAATGTTATTCGAAACATTAACAAATCCAAAGA	2495
Db	5518	TAAACACAAAATCATTCGATTAATAATATTTTAAAAATATATATTAATTAATCAATCA	5455
Qy	2496	TTTTCAAT 2503	
Db	5458	AAATCAAT 5451	

```

RESULT 38
Sequence 70. Application US/10204708
Patent No. 6677731
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PLEMBROCK, Christian
TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
TITLE OF INVENTION: by Assessing DNA Methylation
FILE REFERENCE: 5013.1012
CURRENT APPLICATION NUMBER: US/10/204,708
CURRENT FILING DATE: 2003-05-06
PRIOR APPLICATION NUMBER: PCT/EP01/03971
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: DE 10019058.8
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: DE 10019173.8
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 98
SEQ ID NO 70
LENGTH: 6040
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURES:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-70

```

	Query Match	Similarity	Score	DB	Length
Best local	629	44.7%	4.7e-07		6040
Matches	629	Conservative	0	Mismatches	757
				Indels	20
				Gaps	7
Cy	1107	AACTTGCGATCATGATGAGAAACAATAAGATAGCTTAAATAATATATACATC			1166
Db	3193	AAACGATTAACCTACAGCCCTATATATCCACACTTAAAAATATATAAAAAATATCGC			3134
Oy	1167	AAGATATTATATAGCTACAGATTTGAGCTGCTATTAAAGAACATAATTGCATGTG			1228

Db	3133	AAATCAAAAAATCGAAACCATCTTAACCAACATATATAAAACCTCGTCTCTACTAAAAATA	3074
Qy	1227	GAATATTCGCTGGAAAAATTTAGTAAATAAGCTGTGCTACTGTTATACTAAAGCA	1286
Db	3073	CAAAATTTAACCAACGTAATTAATACAGCGCTTAATATCCCACTACTCGAA-AAACTAAA	3015
Qy	1287	TGAATGAAATTTAAACACTATATATATAGCAAAAAATTCATTCAATGATCTTGACATTTAA	1346
Db	3014	ACAAAAAATTACTTTAAACCCAAAAAACAATAATTAACAATTAATATATACACTA	2955
Qy	1347	TTGGTCCATCACTGCTCTGGAGATCTAATATTAACATGTACACGACATTCACGCAATA	1406
Db	2954	CATCCACCTTAACACACAAAAAATCCCATCTCAAAAAAATAATCATTTAAAT	2895
Qy	1407	TGGCTTTTGA ^g CTAGAAATATAGAAAAAGTAGAAATATTAATATCATTAATAGATCAACAC	1466
Db	2894	ACAATATATCAAAATTTTATTAATCATTAACCATTTAAATTA ^g CTCCCAATAAATCACTAAATTT	2835
Qy	1467	TAAAGCTTGTGAGAGAACACGACTGTGTA ^g AAAAACACAGATATTCATTGGCAAAAAACCTA	1536
Db	2834	AAATACGAATCTAAAAATCTAAAAACATATTCATTTCTTAACTACTAAATTTACCTT	2775
Qy	1537	ATGTAGAACTACCAATTTGCATATCTATTTACAAATTTATTACATGAAATATATCACTAG	1586
Db	2774	CTTAAAACTCACTCTGTAATCTACAAACAATATACCTTTAAATCCCAACATCACAA	2715
Qy	1587	ATAAAGCATATCAAAATATATCTATAGTGTATCATTCCTTGATTAATTCACACACATG	1646
Db	2714	AATAAAAACACCTACCAATTTTTCGGTTTCCACCTATATCTAAATTTATATCTTAT	2655
Qy	1647	ATAAATAAACGGTCAAAAAATATAGAAATACACGTTCCACACATATATTA ^g AAAAAGCAGTCA	1706
Db	2654	ATTTATATAAAAAACAATACATTAATCTTATTTCACTCTTATTAATATATTTCTATCG	2595
Qy	1707	ATACCCATTCGTTACATTCACCACTAGATGTAAATTTATACATACAAACTTAAGTATTA	1766
Db	2584	ATTAACACATATAATTTTGGAAAACTTAACATTTAAATTAATTAATTAATTAATTAATTA	2535
Qy	1767	CTTAAAAATTCCTAGCTAATTAATATATAACAATTCCTATCTCTGATCTCAATTAACCAAT	1826
Db	2534	AATCTACTAAATATATATACGAAAAACAATCAATCTCTCATATTTCACTTCTTCACCTAT	2475
Qy	1827	TGTTATCAGATTAATAAATAAGT-CAATTAACCTTTACTATTTATTAATTAATTAATTAAT	1885
Db	2474	AAACATCTATTCAAAAATTTACTATTAACCTTAATTTCCATCTCTAATTAATTAATTTT	2415
Qy	1886	TGCACACTTTTCACTACATCATG- ----TATATTAAGAACATTAATTTACATTAATA	1939
Db	2414	AAAAAAACCTAATATACATAAAACCTAATTAATTAATTAATTAATTAATTAATTAATTA	2355
Qy	1940	ACTATATAAAGAACATACATATATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	1959
Db	2354	ACAAACATTAACAAATTAATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTA	2295
Qy	2000	TTACATATGATGT- - - - -CCTGACATATATAATTTATATCAATCTTGACATATTTCAACAATA	2056
Db	2294	TTTATCAATCTTAACAATAAATATATAAAAAAATTTATCTTAACTTAACATAGAAATTA	2235
Qy	2057	AATATGATTAACCTTTTTTTTATGCGCATTAAGTCTTAATTAATTTGCTAATTAACCAA	2116
Db	2234	AAAAATTCACATTCGATTTCTTA ^g AAAAAATATATTAATTAATTAATTAATTAATTAATTA	2175
Qy	2117	AAGGTTAAATACATCTCTCGCAACAATATGAAGAAGATACAAATCAAGCTCAATAATAT	2176
Db	2174	-----TCTCACTACCATCTTCCACCAAAATTAATAATTTGCTCCAAAAAATAATTAAT	2120
Qy	2177	TAAACAATTTCTATATTAACACACTACTCTAAACATATAGAACCAATATGATATCATATCC	2236
Db	2119	TATTAATTAATTAATTTCTTAACAACATTTGTATTAACATTAACATTAATTTTTCACCAT	2060
Qy	2237	TAAATTAAGATATCTCTACAGCATGTATCAAAATTAATAAGAAAAATCGTATGTATATTTA	2296

2231 CATACCTAGTATATATCCCTACAGCATGTATCAATAATAAGAAAAATCGTTATGTA 2290
2235 TATTTTATTTCTTAAATATTTCTCCCTAAATATCCAAAAATATAATTAATTTT 2236
2231 TATTTATACATTAACAAATTTCCAACTTCAACACAAACAACTCAAGTCTATTATACA 2350
2235 CTCTTATACATCTCTATATAATATAATATAACCAATTAATTAATTAATTAATTAAT 2176
2351 AATTAAGATATATATATATCTTCAACCAACATAGTTATCAATCTATATATACATCA 2410
2175 CAATATATATTTATATATATCTTAAATTAATTAATTAATTAATTAATTAATTAAT 2116
2411 TCCCTTTTAAAGATGACTTCAATTCATCATTAATTAATTAATTAATTAATTAAT 2456
2115 CCATTTATATATAAAACAAACGATTAATTAATTAATTAATTAATTAATTAATTAAT 2070

RESULT 40
US-10-204-708-29/c
Sequence 29, Application US/10204708
Patent No. 6677731
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPERBROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
TITLE OF INVENTION: By Assessing DNA Methylation
FILE REFERENCE: 5013.1012
CURRENT APPLICATION NUMBER: US/10/204.708
CURRENT FILING DATE: 2003-05-06
PRIOR APPLICATION NUMBER: PCT/EP01/03971
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: DE 10019058.8
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: DE 10019173.8
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 98
SEQ ID NO 29
LENGTH: 5666
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-29

Query Match 1.8%; Score 84.6; DB 4; Length 5666;
Best Local Similarity 44.6%; Pred. No. 5.1e-07;
Matches 1051; Conservative 0; Mismatches 1254; Indels 54; Gaps 16;

149 CTCGCACATTGATATAGAGCAACCATTTTCCAAATCCCAATTTTAACATGCAACA 208
2433 CTCGAAAAAATAAATTTCAATCCCTTTTAAATTAATTAATTAATTTTCAAAACA 2374
209 GATTATTTATGTAAACAGAACCCCTGTACTTATGACAGAACTTAATTAATTAATCA 268
2373 TATTTCCAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2314
269 GGTCTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 328
2313 ATTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTCTCCTA 2260
339 TTTTCCA-ATTATTTACATATCATAGAGTTCTTTATGACATATATTAATTAATCAAT 387
2259 TATTTACAAATTTATCTATCTATTAATTAATTAATTAATTAATTAATTAATTAATTA 2200
388 CAAAACAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 447
2199 ACAAACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2140

448 ACTACAAATATATATGAGACCTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 507
2139 TAAATATATACACCTCCACACTTTTATTTTATTAATTAATTAATTAATTAATTAATTA 2080
508 TCAATTTTGCAATATATACATCTGACATGATTAACATTAATTAATTAATTAATTAATTA 567
2079 TACATTCATTTATATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2020
568 TGAATATATATATATATATATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 627
2019 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1960
628 AAAAAATGAATCAGATTTTGTGACAGATCAATTTGGCAGACAAATGCAATTTGCACT 687
1959 CCTACATACACATTTCTTTAAATCTTCAATTAATTTTAAATTAATTAATTAATTAATTA 1900
688 GTACGACATGGTATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 746
1899 ACCAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1840
747 TAAACACTTACGAAAAATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 806
1839 CTATCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1780
807 ATGCAACGACAAATATATATATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 862
1779 TTTCTTATTTTATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1720
863 ACTATTTCTACACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 921
1719 ACCACCTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1660
922 GTGTAAATATCTCCAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 981
1659 TTTACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1600
982 TCCCACTGAATATGACAGAAATTTTACAAATTAATTAATTAATTAATTAATTAATTAAT 1034
1599 ATTTACAAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1540
1035 CTGGTCAAGTTTGTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1094
1539 TATCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1480
1095 GTGATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1154
1479 TAAACAAACAACTACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1420
1155 TAAATATCAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1213
1419 TAAACAAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1360
1214 GCATATGATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1273
1359 TACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1300
1274 ATTAAGTAAAGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1333
1299 AATACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1240
1334 CTGATATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1380
1239 TATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1180
1381 ATGATACAGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1439
1179 ATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1120
1440 ATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1499
1119 TACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1060

QY 1500 CACTGATATCATAGCAAAAAAAGTTAAATGAGAACTACCAATTTGCAATATCTATTTACA 1559
DB 1059 GCACCTTCCCTGGCCGCACTACATCGCA-ACCTAATATACAAAAGAAATCTCATCTCAA 1001
QY 1560 ATTATATACATGAGAAATATATACATGATTAAGCCATATCAACATATTAATCTTAGTCTA 1619
DB 1000 AAAACCAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 943
QY 1620 TCATCTTTGATTAATTTCAACCAATGATTAATTAACGGTCAAAAATATAGATACACAG 1679
DB 942 CTATATATATTTATTAATAAAAGTCACTTCACTTACACATATTAATAACAAATATATCG 883
QY 1680 TTCACCACTATATATAAAGAGTCAATATCCATGCTTAACATCCAACTAAGTATGTA 1739
DB 882 TATACCAATTAATTAATTTCTATCAATTAACAAATCTATATACAACTATATCTTATTA 823
QY 1740 CA-TATTAATATACAAACTTAAGTATTAATTAATTTCTAGCTAATTAATTAACAAA 1798
DB 822 AACGCAATACATATATTTTAATTAATTTCAATTTTATTAATTTTAATTAATTAACATTA 763
QY 1799 TTCTATCTCTATCTCAATTAACCAATTTGTTATGAGATTAATTAATTAAGTCAATTAAC 1858
DB 762 TACTTACCACTATATACCAATTAATTAAGTATTAATTAACCAATTAATTAATTAATTA 703
QY 1859 TTTTACTTATTAATAATTAATTAATTTGCAACATTTTCTACATCATGATATTAATG 1918
DB 702 TATTAACCTTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 643
QY 1919 AACATATATTTTACATTAATAATTAATTAATAAAGAAACCAATGATATTAATTA--AATTA 1975
DB 642 TCTAATTTTATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 583
QY 1976 AAAACACCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2035
DB 582 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 523
QY 2036 CTGACATATTTTCAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2094
DB 522 ATTTATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 463
QY 2095 ATTAATTAATTTCTAA-----TTACCAAAAAGTTAAATTAACATTAATTAATTAATTA 2144
DB 462 AAACCACTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 403
QY 2145 TATGAAGAAGATACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2204
DB 402 ACTAAACCCAAAACCAAAAACCAAAAACCAAAAACCAAAAACCAAAAACCAAAAACCAAT 343
QY 2205 TCTAAACATAGACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2264
DB 342 ATTTAAATAATTAATAATTAATAATTAATAATTAATAATTAATAATTAATAATTAATAAT 283
QY 2265 AAATAATAATAATAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2324
DB 282 CAATAATAATCAATAATTTCTAACAATAATAATAATAATAATAATAATAATAATAATAAT 223
QY 2325 ACATAAATCTCA-AGTTCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2383
DB 222 ACCAATATCAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 163
QY 2384 AGTTATCAATTTCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2443
DB 162 ATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 103
QY 2444 ATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2462
DB 102 ATTAATTTCTCAAAAATCA 84

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2004, 18:26:43 ; Search time 1530 Seconds

(without alignments)
11633.871 Million cell updates/sec

Title: US-10-081-051-52

Perfect score: 4750

Sequence: 1 gatccgtcattactatgag.....ccacattgactatgac 4750

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2470430 seqs, 1873665578 residues

Total number of hits satisfying chosen parameters: 4240860

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4750	100.0	4750	US-10-081-051-52	Sequence 52, Appl
2	1416	29.8	1416	US-10-081-051-56	Sequence 56, Appl
3	984	20.7	984	US-10-081-051-54	Sequence 54, Appl
4	597	12.6	597	US-10-081-051-57	Sequence 57, Appl
5	371	7.8	371	US-10-081-051-53	Sequence 53, Appl
6	269.2	5.7	1371	US-10-010-160-5	Sequence 53, Appl
7	207.8	4.4	3673778	US-10-312-841-1	Sequence 1, Appl
8	187.6	3.9	1017	US-10-282-122A-16370	Sequence 16370, A
9	175.2	3.7	1029	US-10-282-122A-17353	Sequence 17353, A
10	170.6	3.6	996	US-10-282-122A-35500	Sequence 35500, A
11	168.6	3.5	989	US-09-815-242-8974	Sequence 8974, Ap
12	168.4	3.5	999	US-10-282-122A-15395	Sequence 15395, A
13	165.4	3.5	999	US-10-282-122A-34158	Sequence 34158, A
14	155.2	3.3	7814	US-10-221-714A-252	Sequence 252, App
15	155	3.3	1065	US-10-282-122A-21770	Sequence 21770, A

C 16	155	3.3	5641	14	US-10-311-455-1369	Sequence 1369, Ap
C 17	154.2	3.2	918	8	US-08-781-986A-513	Sequence 513, App
C 18	154.2	3.2	918	12	US-10-329-624-513	Sequence 513, App
C 19	153	3.2	15548	14	US-10-311-455-2128	Sequence 2128, Ap
C 20	148.8	3.1	6314	14	US-10-240-452-14	Sequence 14, Appl
C 21	148.4	3.1	1020	12	US-10-282-122A-6338	Sequence 6338, Ap
C 22	148.4	3.1	1023	9	US-09-815-242-6755	Sequence 6755, Ap
C 23	147.2	3.1	1026	9	US-09-815-242-3869	Sequence 3869, Ap
C 24	147.2	3.1	6132	14	US-10-311-455-836	Sequence 836, App
C 25	146.2	3.1	319630	15	US-10-398-221-7	Sequence 7, Appl
C 26	146.2	3.1	3013208	15	US-10-398-221-2058	Sequence 2058, Ap
C 27	145.2	3.1	11745	14	US-10-240-453-206	Sequence 206, App
C 28	145.2	3.0	7597	14	US-10-311-455-886	Sequence 986, App
C 29	141.6	3.0	3673778	14	US-10-312-841-2	Sequence 2, Appl
C 30	139.8	2.9	2791	15	US-10-398-221-3618	Sequence 3618, Ap
C 31	139	2.9	6175	14	US-10-311-455-1280	Sequence 1280, Ap
C 32	137	2.9	18154	14	US-10-311-455-327	Sequence 227, App
C 33	136.8	2.9	1008	9	US-09-815-242-6968	Sequence 6968, Ap
C 34	136.8	2.9	1008	12	US-10-282-122A-22052	Sequence 22052, A
C 35	136.8	2.9	1830121	14	US-10-329-960-1	Sequence 1, Appl
C 36	136.8	2.9	1830121	15	US-10-329-960-1	Sequence 1, Appl
C 37	136.2	2.9	11691	14	US-10-311-455-2213	Sequence 2213, Ap
C 38	136	2.9	17594	14	US-10-311-455-1999	Sequence 1999, Ap
C 39	135.8	2.9	6831	14	US-10-311-455-1460	Sequence 1460, Ap
C 40	135.4	2.9	11691	14	US-10-311-455-2214	Sequence 2214, Ap
C 41	134.8	2.8	15387	14	US-10-311-455-157	Sequence 157, App
C 42	134.4	2.8	6106	12	US-10-221-714A-151	Sequence 151, App
C 43	134.4	2.8	6106	14	US-10-311-455-1445	Sequence 1445, Ap
C 44	134.4	2.8	6106	16	US-10-257-166-113	Sequence 113, App
C 45	134	2.8	7047	14	US-10-240-453-260	Sequence 260, App

ALIGNMENTS

RESULT 1
US-10-081-051-52
Sequence 52, Application US/10081051
Publication No. US2003004422A1
GENERAL INFORMATION:
APPLICANT: Barbet, Anthony F.
APPLICANT: Whitmore, William W.
APPLICANT: Kamper, Sondra M.
APPLICANT: Simbi, Bigboy H.
APPLICANT: Ganta, Roman R.
APPLICANT: Moreland, Annie L.
APPLICANT: Mwangi, Duncan M.
APPLICANT: McGuire, Travis C.
APPLICANT: Mahan, Sumant M.
TITLE OF INVENTION: Ehrlichia Ruminantium Polypeptides, Antigens, Polynucleotides, a
FILE OF INVENTION: Methods of Use
FILE REFERENCE: US-299XCI
CURRENT APPLICATION NUMBER: US/10/081, 051
CURRENT FILING DATE: 2002-02-20
PRIOR APPLICATION NUMBER: US 60/269, 944
PRIOR FILING DATE: 2001-02-20
NUMBER OF SEQ ID NOS: 117
SOFTWARE: PatentIn version 3.1
SEQ ID NO 52
LENGTH: 4750
TYPE: DNA
ORGANISM: Ehrlichia ruminantium (formerly Cowdria ruminantium)
US-10-081-051-52

Query Match 100.0%; Score 4750; DB 14; Length 4750;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4750; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATCCGCTAATCTTACGTAATGTTTAAAGTGGGCTGATATAAATAACT 60
DB 1 GATCCGCTAATCTTACGTAATGTTTAAAGTGGGCTGATATAAATAACT 60
QY 61 GCTTGAATGCTATATATAGTACATACGTAATGATTAAGAAAAATTCTCAAGCAGAC 120

```
Db 61 GCTTACCTCTATATAGTACATACGTATGATGAAAGAAAAATTTCTCAAGAGAC 120
Qy 121 AGTACCTTGCTATGTTAGTCTGCTTCACTCCACATTTGATCTAGAGAAACATTTT 180
Db 121 AGTACCTTGCTATGTTAGTCTGCTTCACTCCACATTTGATCTAGAGAAACATTTT 180
Qy 181 CTCAATCTGCAATTTTAAAGTACGACAAAGATTTATATGTAACAGAACCTGTATCT 240
Db 181 CTCAATCTGCAATTTTAAAGTACGACAAAGATTTATATGTAACAGAACCTGTATCT 240
Qy 241 TTATGCAAGAACTCTATTAATTTGCAAGGTCTCTTACATAGAAACATGTTATCAA 300
Db 241 TTATGCAAGAACTCTATTAATTTGCAAGGTCTCTTACATAGAAACATGTTATCAA 300
Qy 301 GAATACCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Db 301 GAATACCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Qy 361 TTATTCGACATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 420
Db 361 TTATTCGACATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 420
Qy 421 AAAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Db 421 AAAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Qy 481 TATATTCGACATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 540
Db 481 TATATTCGACATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 540
Qy 541 TAAATATATATATATATATATATATATATATATATATATATATATATATATAT 600
Db 541 TAAATATATATATATATATATATATATATATATATATATATATATATATATAT 600
Qy 601 GACTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Db 601 GACTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Qy 661 ATTTGACGACATATATATATATATATATATATATATATATATATATATATAT 720
Db 661 ATTTGACGACATATATATATATATATATATATATATATATATATATATATAT 720
Qy 721 ACCTGATCATAGAAATATATATATATATATATATATATATATATATATATAT 780
Db 721 ACCTGATCATAGAAATATATATATATATATATATATATATATATATATATAT 780
Qy 781 CACATATCATATATATATATATATATATATATATATATATATATATATATAT 840
Db 781 CACATATCATATATATATATATATATATATATATATATATATATATATATAT 840
Qy 841 CAACAATATATATATATATATATATATATATATATATATATATATATATAT 900
Db 841 CAACAATATATATATATATATATATATATATATATATATATATATATATAT 900
Qy 901 AATATCAACAAACAGATATATATATATATATATATATATATATATATATATAT 960
Db 901 AATATCAACAAACAGATATATATATATATATATATATATATATATATATATAT 960
Qy 961 CGAAATATATATATATATATATATATATATATATATATATATATATATATAT 1020
Db 961 CGAAATATATATATATATATATATATATATATATATATATATATATATATAT 1020
Qy 1021 AATTTTATATATATATATATATATATATATATATATATATATATATATATAT 1080
Db 1021 AATTTTATATATATATATATATATATATATATATATATATATATATATATAT 1080
Qy 1081 TATATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
Db 1081 TATATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
Qy 1141 TGATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
Db 1141 TGATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200

Db 1141 TGATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
Qy 1201 AAGAAACATATATATATATATATATATATATATATATATATATATATATAT 1260
Db 1201 AAGAAACATATATATATATATATATATATATATATATATATATATATATAT 1260
Qy 1261 TGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
Db 1261 TGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
Qy 1321 TCATTCATATATATATATATATATATATATATATATATATATATATATATAT 1380
Db 1321 TCATTCATATATATATATATATATATATATATATATATATATATATATATAT 1380
Qy 1381 ATGTCACACAGAACTTACGCAATATATATATATATATATATATATATATATAT 1440
Db 1381 ATGTCACACAGAACTTACGCAATATATATATATATATATATATATATATATAT 1440
Qy 1441 TATAATATATATATATATATATATATATATATATATATATATATATATATAT 1500
Db 1441 TATAATATATATATATATATATATATATATATATATATATATATATATATAT 1500
Qy 1501 ACTGATATATATATATATATATATATATATATATATATATATATATATATAT 1560
Db 1501 ACTGATATATATATATATATATATATATATATATATATATATATATATATAT 1560
Qy 1561 TTTATATATATATATATATATATATATATATATATATATATATATATATATAT 1620
Db 1561 TTTATATATATATATATATATATATATATATATATATATATATATATATATAT 1620
Qy 1621 CATCTTATATATATATATATATATATATATATATATATATATATATATATAT 1680
Db 1621 CATCTTATATATATATATATATATATATATATATATATATATATATATATAT 1680
Qy 1681 TCACACATATATATATATATATATATATATATATATATATATATATATATAT 1740
Db 1681 TCACACATATATATATATATATATATATATATATATATATATATATATATAT 1740
Qy 1741 ATATATATATATATATATATATATATATATATATATATATATATATATATAT 1800
Db 1741 ATATATATATATATATATATATATATATATATATATATATATATATATATAT 1800
Qy 1801 CCTATCTCTATATATATATATATATATATATATATATATATATATATATATAT 1860
Db 1801 CCTATCTCTATATATATATATATATATATATATATATATATATATATATATAT 1860
Qy 1861 TTAATATATATATATATATATATATATATATATATATATATATATATATATAT 1920
Db 1861 TTAATATATATATATATATATATATATATATATATATATATATATATATATAT 1920
Qy 1921 CATATATATATATATATATATATATATATATATATATATATATATATATATAT 1980
Db 1921 CATATATATATATATATATATATATATATATATATATATATATATATATATAT 1980
Qy 1981 ACCTATATATATATATATATATATATATATATATATATATATATATATATATAT 2040
Db 1981 ACCTATATATATATATATATATATATATATATATATATATATATATATATATAT 2040
Qy 2041 CATATATATATATATATATATATATATATATATATATATATATATATATATAT 2100
Db 2041 CATATATATATATATATATATATATATATATATATATATATATATATATATAT 2100
Qy 2101 TTGCTATATATATATATATATATATATATATATATATATATATATATATATAT 2160
Db 2101 TTGCTATATATATATATATATATATATATATATATATATATATATATATATAT 2160
Qy 2161 TACAGCTCAATATATATATATATATATATATATATATATATATATATATATATAT 2220
Db 2161 TACAGCTCAATATATATATATATATATATATATATATATATATATATATATATAT 2220
Qy 2221 AATATATATATATATATATATATATATATATATATATATATATATATATATATAT 2280
Db 2221 AATATATATATATATATATATATATATATATATATATATATATATATATATATAT 2280
```

OY	228	TGTTATGTATATTATATCAATTAACAAATTTCCAACTTGCAGAACGAACAACGCCAAGTC	2340
Db	2281	TGCTTAGTATATTATTATTCAAATTCACAAATTTCCAACTTGCAGAACGAACAACGCCAAGTC	2340
OY	2341	TATTATTAACAATTCAGAATTTATATATCTTCACCAACACATGGTATTCGAATCTTATA	2400
Db	2341	TATTATTAACAATTCAGAATTTATATATCTTCACCAACACATGGTATTCGAATCTTATA	2400
OY	2401	ATACATCATATCCTTTTAAAAAGATGACTTCACAAATTCATACATATTAATATATATCATG	2460
Db	2401	ATACATCATATCCTTTTAAAAAGATGACTTCACAAATTCATACATATTAATATATATCATG	2460
OY	2461	CATGTTTTATTACGAACATTACACAAATTCAGAAATTTCAAATTAACGTTGTAATGCTGA	2520
Db	2461	CATGTTTTATTACGAACATTACACAAATTCAGAAATTTCAAATTAACGTTGTAATGCTGA	2520
OY	2521	ACGTTCCATTCGACAAATTCAGAGCTTCGATATCATATTAACCTCCAAAACAGATTAAGTG	2580
Db	2521	ACGTTCCATTCGACAAATTCAGAGCTTCGATATCATATTAACCTCCAAAACAGATTAAGTG	2580
OY	2581	AGTTTTTAAATACGTCCTTTCAAATCTTCCACGAGCTTTAGTATGGTATACAGAAATAC	2640
Db	2581	AGTTTTTAAATACGTCCTTTCAAATCTTCCACGAGCTTTAGTATGGTATACAGAAATAC	2640
OY	2641	TTTACGACTTAAACATCATTAATAGGGGAATTAGATCTATATCTATCTATGGTAAATCTTT	2700
Db	2641	TTTACGACTTAAACATCATTAATAGGGGAATTAGATCTATATCTATCTATGGTAAATCTTT	2700
OY	2701	TGCTGTATATCATCTTCCTTAGAGATTTCAATATTAATPAAATCCATTCATACATTAAGTAA	2760
Db	2701	TGCTGTATATCATCTTCCTTAGAGATTTCAATATTAATPAAATCCATTCATACATTAAGTAA	2760
OY	2761	TTGCTTTAGTATACCTGGCCATTCATATGATCGCATTTGCTATTAAGTTCAATCACTTAA	2820
Db	2761	TTGCTTTAGTATACCTGGCCATTCATATGATCGCATTTGCTATTAAGTTCAATCACTTAA	2820
OY	2821	TACATGAGTACACAAACCTATTTTTTTAGAGATGCTATTCATPAAATACGTAACAATTC	2880
Db	2821	TACATGAGTACACAAACCTATTTTTTTAGAGATGCTATTCATPAAATACGTAACAATTC	2880
OY	2881	CGATATATCTGTACAAATTTCTACTPAAAGA CGATACCTTAATTTGGAAGGACATTTAATCT	2940
Db	2881	CGATATATCTGTACAAATTTCTACTPAAAGA CGATACCTTAATTTGGAAGGACATTTAATCT	2940
OY	2941	ATATATPAAATCTCTCACAAACCTACACGCTTTTACTTCACTTCAATATCTTTGGAAAGA	3000
Db	2941	ATATATPAAATCTCTCACAAACCTACACGCTTTTACTTCACTTCAATATCTTTGGAAAGA	3000
OY	3001	AGACACATATATTTCTACATCTATATCTAACAGAAATCTTAATTTTCCCATATATATTTT	3060
Db	3001	AGACACATATATTTCTACATCTATATCTAACAGAAATCTTAATTTTCCCATATATATTTT	3060
OY	3061	TCCCTCCCTGATAATCTGAGTAATCTTAATTTGGGATACGTAATCTGTTACTTC	3120
Db	3061	TCCCTCCCTGATAATCTGAGTAATCTTAATTTGGGATACGTAATCTGTTACTTC	3120
OY	3121	ATCTATPAAATAGAGTACCAGATTTGCTGCTCATATATTTCCAATATAGAGAGGTACTCT	3180
Db	3121	ATCTATPAAATAGAGTACCAGATTTGCTGCTCATATATTTCCAATATAGAGAGGTACTCT	3180
OY	3181	ATGAGACAAATATATTTTAACTTTTCCCTACCTACCAAATATATTAACCAAGTAAATATTAGC	3240
Db	3181	ATGAGACAAATATATTTTAACTTTTCCCTACCTACCAAATATATTAACCAAGTAAATATTAGC	3240
OY	3241	TGTTAGCATATAGTACATAGATATTAATAGAGATATCATACCCCTTGATTTTTATG	3300
Db	3241	TGTTAGCATATAGTACATAGATATTAATAGAGATATCATACCCCTTGATTTTTATG	3300
OY	3301	TATTAGCTACCAATCTCTTCCAAACCTGGGGAACCAATATAGATATAGACT	3360
Db	3301	TATTAGCTACCAATCTCTTCCAAACCTGGGGAACCAATATAGATATAGACT	3360

QY	3361	CGANGTGTACTGCTTTATTAACATACCTCTCAAAATTCAGTAAACAGGGAGTTAC	3420
Db	3361	CGATGTAGTACTGCTTTATTAACATACCTCTCAAAATTCAGTAAACAGGGAGTTAC	3420
QY	3421	GACTATTTCAATACTCCAAATGCTGATTTCACTCATCTTTCTGACGTAATCTAC	3460
Db	3421	GACTATTTCAATAATCCCAATGCTGATTTCACTCATCTTTCTGACGTAATCTAC	3460
QY	3461	AGACTCTATAGCTCTCTTACACACTAATCTTAATCTCTCTCTGTAAAGGCTTTCTAT	3540
Db	3461	AGACTCTATAGCTCTCTTACACACTAATCTTAATCTCTCTCTGTAAAGGCTTTCTAT	3540
QY	3541	ATTAATCATMAGCACCCATATGACAGACTTTACGCACTGGCAATATTACATGCCACT	3600
Db	3541	ATTAATCATMAGCACCCATATGACAGACTTTACGCACTGGCAATATTACATGCCACT	3600
QY	3601	AATATATAATAACAGGCAATATAGATATACCTTTTAAAGCTTTCCAGTACATTAATCC	3660
Db	3601	AATATATAATAACAGGCAATATAGATATACCTTTTAAAGCTTTCCAGTACATTAATCC	3660
QY	3661	ATCATATACAGATCTCTTAACCATATATCCAAATATACAAATCAGGCTTTTTCATA	3720
Db	3661	ATCATATACAGATCTCTTAACCATATATCCAAATATACAAATCAGGCTTTTTCATA	3720
QY	3721	AGCCATCTGATCGCGGATAAACCATCACTGCTAATTTTGTGACATTAATATCATCACT	3780
Db	3721	AGCCATCTGATCGCGGATAAACCATCACTGCTAATTTTGTGACATTAATATCATCACT	3780
QY	3781	TAAATATCTTTTATTTGATTTCTGATATCACTTCAATCAACAACTAATACTTACAGA	3840
Db	3781	TAAATATCTTTTATTTGATTTCTGATATCACTTCAATCAACAACTAATACTTACAGA	3840
QY	3841	AATATACATCTTTCCTTGACATTTCAAAATCCTGTGCATATATTAATCTACATATCTAA	3900
Db	3841	AATATACATCTTTCCTTGACATTTCAAAATCCTGTGCATATATTAATCTACATATCTAA	3900
QY	3901	TGAAATATCTGAAACCAATATGACTCCTTTAAAGTACACTCAATATCTACACACACA	3960
Db	3901	TGAAATATCTGAAACCAATATGACTCCTTTAAAGTACACTCAATATCTACACACACA	3960
QY	3961	CTAATGATCAACCAACCAATATTAATTAATCTATTTTAACTTAATTAATTAACATAT	4020
Db	3961	CTAATGATCAACCAACCAATATTAATTAATCTATTTTAACTTAATTAATTAACATAT	4020
QY	4021	TATATAAATTTGTTACTATTTATCTAATTTATACAGCATATAGTTTCTATATATCCAA	4080
Db	4021	TATATAAATTTGTTACTATTTATCTAATTTATACAGCATATAGTTTCTATATATCCAA	4080
QY	4081	AAGAAATATATAAACACTTATATATAAACAATTATCATCACTTTGTAACCTATAGTACT	4140
Db	4081	AAGAAATATATAAACACTTATATATAAACAATTATCATCACTTTGTAACCTATAGTACT	4140
QY	4141	AATATATTAATTAATTTTATTTATTTATTAATTAACCTTCCGTGAAACAGCATATCTATAGTT	4200
Db	4141	AATATATTAATTAATTTTATTTATTTATTAATTAACCTTCCGTGAAACAGCATATCTATAGTT	4200
QY	4201	TCCATATCTAATAAATCTCACTTGAAGCTTTATAGATATTCATTAATTTACGGCAACAGCAT	4260
Db	4201	TCCATATCTAATAAATCTCACTTGAAGCTTTATAGATATTCATTAATTTACGGCAACAGCAT	4260
QY	4261	AAGAAATCATTAATTTAGTAAATTTTATATTTAAACCACTAGCAAAATTAATATCCACCAAT	4320
Db	4261	AAGAAATCATTAATTTAGTAAATTTTATATTTAAACCACTAGCAAAATTAATATCCACCAAT	4320
QY	4321	GACTATCAACATATGTATAGATGAAACCAATACCAACCAACAAAGATATTATCAATTGACTA	4380
Db	4321	GACTATCAACATATGTATAGATGAAACCAATACCAACCAACAAAGATATTATCAATTGACTA	4380
QY	4381	CTTACAGTAATAAATCTAGATGAGTTTAAGCACACATCTCCATATGTAAGTCAAAATCAT	4440
Db	4381	CTTACAGTAATAAATCTAGATGAGTTTAAGCACACATCTCCATATGTAAGTCAAAATCAT	4440
QY	4441	CCCGTATATTAACATTAATTCACAATAATATATGTCCAAAACCTATCAATGATAGAAATTTTGT	4500

Db	12	TAATACTGCTTTCAAAATCTTCACAGAGCTTAAAGTAATGGTACAGAAATTAACTTAGC	180
QY	2647	ACTTAAACACATCATTAATAAGCGAAATTAGTACTATATCTACTGGTAAATCTTTGCGT	2766
Db	181	ACTTAAACACATCATTAATAAGCGAAATTAGTACTATATCTACTGGTAAATCTTTGCGT	240
QY	2707	AATCATCTCCTTAGAGAAATTTCATTAATTAATTCATTTCTATATACATTTAGTAAATGCT	2766
Db	241	AATCATCTCCTTAGAGAAATTTCATTAATTAATTCATTTCTATATACATTTAGTAAATGCT	300
QY	2767	TAAATTACCTGGCCATTCAATAGCTGCATTTGGTATTAAGCTTCAGACCTTAATACATG	2826
Db	301	TAAATTACCTGGCCATTCAATAGCTGCATTTGGTATTAAGCTTCAGACCTTAATACATG	360
QY	2827	AGTACACAAACCTAATTTTTTTTACAGATGCTATTCTATTAATAATCCTACACAAATTCGGTAT	2886
Db	361	AGTACACAAACCTAATTTTTTTTACAGATGCTATTCTATTAATAATCCTACACAAATTCGGTAT	420
QY	2887	ATCTGTACAAATATCTCTCTAAGAGCGGTACTCTTAATTTGAAGACATTTATCTATATA	2946
Db	421	ATCTGTACAAATATCTCTCTAAGAGCGGTACTCTTAATTTGAAGACATTTATCTATATA	480
QY	2947	TAAATCTCTCAAAACCTACACAGCTTTTACTTCACTTTCAATATCTTTGAGAAGACAC	3006
Db	481	TAAATCTCTCAAAACCTACACAGCTTTTACTTCACTTTCAATATCTTTGAGAAGACAC	540
QY	3007	AATTAATCTCAACATCTAATCTAAGAAATCTTAATCTTTTCCCTATATATTTTCCCTC	3066
Db	541	AATTAATCTCAACATCTAATCTAAGAAATCTTAATCTTTTCCCTATATATTTTCCCTC	600
QY	3067	CTGTAAATTAATCTGAGTAATCTTAATTGCGTATGCTATCGTAATCTGTACTTCACTAT	3126
Db	601	CTGTAAATTAATCTGAGTAATCTTAATTGCGTATGCTATCGTAATCTGTACTTCACTAT	660
QY	3127	AAATPAAGTACCAAGATTTGGTGGCTATTAATTCGAAATGAGAGGTACTCATGAGA	3186
Db	661	AAATPAAGTACCAAGATTTGGTGGCTATTAATTCGAAATGAGAGGTACTCATGAGA	720
QY	3187	CAATATATATTAATCTTTCTCTCACTACCAAAATATTTAACAAGTAATTAATAGCTGGTAG	3246
Db	721	CAATATATATTAATCTTTCTCTCACTACCAAAATATTTAACAAGTAATTAATAGCTGGTAG	780
QY	3247	CATGATGAGAGACATAGATTAATTAATGAGATCATACCCCTGGATTTTTATGATATAG	3306
Db	781	CATGATGAGAGACATAGATTAATTAATGAGATCATACCCCTGGATTTTTATGATATAG	840
QY	3307	CTTNGCAACTACTTCCTTTCCACACCTGGCGAACAGTAATGATAGATGAGCTCGATGT	3366
Db	841	CTTNGCAACTACTTCCTTTCCACACCTGGCGAACAGTAATGATAGATGAGCTCGATGT	900
QY	3367	AGTAGCGCTTATTAATCATACTCTTCCAAATTCGATATAACAGAGGAGTAAACGATAT	3426
Db	901	AGTAGCGCTTATTAATCATACTCTTCCAAATTCGATATAACAGAGGAGTAAACGATAT	960
QY	3427	TTCTATTAATCTCTCAAAATGCTGATTTCAACTCATATTTTCTCTAGTAATCTACAGACTC	3486
Db	961	TTCTATTAATCTCTCAAAATGCTGATTTCAACTCATATTTTCTCTAGTAATCTACAGACTC	1020
QY	3487	TATAGCTCTCTTTCAGACTAACTTAATCTCTCTCTGTAAAGGCTTTTCTATATATATC	3546
Db	1021	TATAGCTCTCTTTCAGACTAACTTAATCTCTCTCTGTAAAGGCTTTTCTATATATATC	1080
QY	3547	ATAAGCACCAATATATGACAGAGCTTTACAGACATGAGGCAATTTACAGATGCCATTAATCAT	3606
Db	1081	ATAAGCACCAATATATGACAGAGCTTTACAGACATGAGGCAATTTACAGATGCCATTAATCAT	1140
QY	3607	AATAACAGGCAATAATAGATACCTTTCTTTAAGCTTTTCCAGTACACTTAATCCATCAAT	3666
Db	1141	AATAACAGGCAATAATAGATACCTTTCTTTAAGCTTTTCCAGTACACTTAATCCATCAAT	1200
QY	3667	ATCAGATCTCTTTAACAATATATGCAATTAATACAACATCAAGGCTCTTTTCATATAGCAT	3726
Db	1201	ATCAGATCTCTTTAACAATATATGCAATTAATACAACATCAAGGCTCTTTTCATATAGCAT	1260

Qy 3727 CTGATCGGATTAACCATCACTGCTAATTAGTGCATATATATCATCTTAATAT 3786
 Db 1261 CTGATCGGATTAACCATCACTGCTAATTAGTGCATATATATCATCTTAATAT 1320
 Qy 3787 ATCTTTATTTAGATTTCTGATATCACTTATCATCAACATTAATCTTCAAGAAATTA 3846
 Db 1321 ATCTTTATTTAGATTTCTGATATCACTTATCATCAACATTAATCTTCAAGAAATTA 1380
 Qy 3847 CAATCTTCTTGGACATTTCAAAATCCTGTGCCAT 3882
 Db 1381 CAATCTTCTTGGACATTTCAAAATCCTGTGCCAT 1416

RESULT 3

US-10-081-051-54
 ; Sequence 54, Application US/10081051
 ; Publication No. US20030044422A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Barbet, Anthony F.
 ; APPLICANT: Whitmire, William M.
 ; APPLICANT: Kamper, Sondra M.
 ; APPLICANT: Simbi, Bigboy H.
 ; APPLICANT: Ganta, Roman R.
 ; APPLICANT: Moreland, Annie L.
 ; APPLICANT: Mwangi, Duncan M.
 ; APPLICANT: McGuire, Travis C.
 ; APPLICANT: Mahan, Suman M.
 ; TITLE OF INVENTION: *Escherichia Ruminantium* Polypeptides, Antigens, Polynucleotides, and
 ; TITLE OF INVENTION: Methods of Use
 ; FILE REFERENCE: US-299XCI
 ; CURRENT APPLICATION NUMBER: US/10/081, 051
 ; CURRENT FILING DATE: 2002-02-20
 ; PRIOR APPLICATION NUMBER: US 60/269,944
 ; PRIOR FILING DATE: 2001-02-20
 ; NUMBER OF SEQ ID NOS: 117
 ; SOFTWARE: Patent version 3.1
 ; SEQ ID NO 54
 ; LENGTH: 984
 ; TYPE: DNA
 ; ORGANISM: *Escherichia ruminantium* (formerly *Cowdria ruminantium*)
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(984)
 ; OTHER INFORMATION: Corresponds to SEQ ID NO:52, nucleotides 632..1615
 ; OTHER INFORMATION: Hypothetical glycerol 3-phosphate dehydrogenase
 ; OTHER INFORMATION: Product = "19hworf2"
 ; US-10-081-051-54

Query Match 20.7%; Score 984; DB 14; Length 984;
 Best Local Similarity 100.0%; Pred. No. 1.3e-146;
 Matches 984; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 632 ATGAAATCACTATTTAGTGCAGATCATTGGACAGCAAGCAATGCAATGGCTGCA 691
 Db 1 ATGAAATCACTATTTAGTGCAGATCATTGGACAGCAAGCAATGCAATGGCTGCA 60
 Qy 692 GCAATGCTATATCACTTATTTGGGAGCTGATCATGAAATATTACATATTAAC 751
 Db 61 GCAATGCTATATCACTTATTTGGGAGCTGATCATGAAATATTACATATTAAC 120
 Qy 752 ACTTACGCAAAATTTAAATATTTACCCACATATCATCTTACGCAACATATATGCA 811
 Db 121 ACTTACGCAAAATTTAAATATTTACCCACATATCATCTTACGCAACATATATGCA 150
 Qy 812 ACCGCAATATATGAGAAATTTATCTGCAACAATATCATGATTTACTTAACTATTCCT 871
 Db 181 ACCGCAATATATGAGAAATTTATCTGCAACAATATCATGATTTACTTAACTATTCCT 240
 Qy 872 ACACAAATATATGAGCAATATGATGACAAATATACAAACAAGCAATATGTTAAATAT 931
 Db 241 ACACAAATATATGAGCAATATGATGACAAATATACAAACAAGCAATATGTTAAATAT 300

Qy 932 ACTCCAAATTAATTTGATGTAAGGATTCGAAATTAATCATCACTCAATTTCCAGTGA 991
 Db 301 ACTCCAAATTAATTTGATGTAAGGATTCGAAATTAATCATCACTCAATTTCCAGTGA 360
 Qy 992 ATAGCAAGAAATTTTCAATATATCAATTTTATACCTCTGGTCCAGTTTCT 1051
 Db 361 ATAGCAAGAAATTTTCAATATATCAATTTTATACCTCTGGTCCAGTTTCT 420
 Qy 1052 AAAGAAATTCGAAACATCTTCTGTGTATAGTACTGCTGTGTATTAAGAACTT 1111
 Db 421 AAAGAAATTCGAAACATCTTCTGTGTATAGTACTGCTGTGTATTAAGAACTT 480
 Qy 1112 GGTGAATCATATATGAAACAAATAGTATAGTATGTTCTTAAATATATACCATCAAGAT 1171
 Db 481 GGTGAATCATATATGAAACAAATAGTATAGTATGTTCTTAAATATATACCATCAAGAT 540
 Qy 1172 ATTAATAGTGTACAGATTTGAGCTGCATTAAGAAACATATTTGCAATTTGCAATGTA 1231
 Db 541 ATTAATAGTGTACAGATTTGAGCTGCATTAAGAAACATATTTGCAATTTGCAATGTA 600
 Qy 1232 ATGCTGTGAAAAAATTTAGTATATATGCTGTTGCTACTGTTATTAAGGCAATGAT 1291
 Db 601 ATGCTGTGAAAAAATTTAGTATATATGCTGTTGCTACTGTTATTAAGGCAATGAT 660
 Qy 1292 GAAATTAACACATATATATAGCAAAAAATCATTCATATAGATCTCATATATTAATGCT 1351
 Db 661 GAAATTAACACATATATATAGCAAAAAATCATTCATATAGATCTCATATATTAATGCT 720
 Qy 1352 CCATCATGCTTGGAGATCTAATATTAACATGTACACAGAAACATTCACGCAATATGCT 1411
 Db 721 CCATCATGCTTGGAGATCTAATATTAACATGTACACAGAAACATTCACGCAATATGCT 780
 Qy 1412 TTGGACATGAATATGAAAAAGTGAATATTAATTAATTAATTAATTAATTAATTAATTA 1471
 Db 781 TTGGACATGAATATGAAAAAGTGAATATTAATTAATTAATTAATTAATTAATTAATTA 840
 Qy 1472 CTGTGTAAGAAACCAAGTACTGTAAACCACTGATATCATTAAGCAAAAACTTAATGTA 1531
 Db 841 CTGTGTAAGAAACCAAGTACTGTAAACCACTGATATCATTAAGCAAAAACTTAATGTA 900
 Qy 1532 GAACATACCAATTTGATATCTATTTTACAAATTTATACATGGAATATATCACTATATTA 1591
 Db 901 GAACATACCAATTTGATATCTATTTTACAAATTTATACATGGAATATATCACTATATTA 960
 Qy 1592 GCCATATCAACATATTTATCTTAG 1615
 Db 961 GCCATATCAACATATTTATCTTAG 984

RESULT 4

US-10-081-051-57
 ; Sequence 57, Application US/10081051
 ; Publication No. US20030044422A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Barbet, Anthony F.
 ; APPLICANT: Whitmire, William M.
 ; APPLICANT: Kamper, Sondra M.
 ; APPLICANT: Simbi, Bigboy H.
 ; APPLICANT: Ganta, Roman R.
 ; APPLICANT: Moreland, Annie L.
 ; APPLICANT: Mwangi, Duncan M.
 ; APPLICANT: McGuire, Travis C.
 ; APPLICANT: Mahan, Suman M.
 ; TITLE OF INVENTION: *Escherichia Ruminantium* Polypeptides, Antigens, Polynucleotides, and
 ; TITLE OF INVENTION: Methods of Use
 ; FILE REFERENCE: US-299XCI
 ; CURRENT APPLICATION NUMBER: US/10/081, 051
 ; CURRENT FILING DATE: 2002-02-20
 ; PRIOR APPLICATION NUMBER: US 60/269,944
 ; PRIOR FILING DATE: 2001-02-20
 ; NUMBER OF SEQ ID NOS: 117
 ; SOFTWARE: Patent version 3.1
 ; SEQ ID NO 57

```

RESULT 5
US-10-081-051-53
; Sequence 53, Application US/10081051
; Publication No. US20030044422A1
; GENERAL INFORMATION:
; APPLICANT: Barbet, Anthony F.
; APPLICANT: Whitmire, William M.
; APPLICANT: Kamper, Sandra M.
; APPLICANT: Simbi, Bigboy H.
; APPLICANT: Ganta, Roman R.
; APPLICANT: Moreland, Annie L.
; APPLICANT: Mwangi, Duncan M.
; APPLICANT: McGuire, Travis C.
; APPLICANT: Mahan, Suman M.
; TITLE OF INVENTION: Ethelchia Ruminantium Polypeptides, Antigens, Polynucleotides, and
; FILE OF INVENTION: Methods of Use
; FILE REFERENCE: UF-299XCI

```

```

RESULT 6 -US-10-010-160-5/c
; Sequence 5, Application US/10010160
; Publication No. US20030103999A1
; GENERAL INFORMATION:
; APPLICANT: Rosey, Everett L.
; APPLICANT: Strugnell, Richard A.
; APPLICANT: Good, Robert T.
; APPLICANT: King, Kendall W.
; TITLE OF INVENTION: NOVEL THERAPEUTIC COMPOSITIONS FOR
; TREATING INFECTION BY LAWSONIA SPP
; FILE REFERENCE: DAV1101.001aUS
; CURRENT APPLICATION NUMBER: US/10/010.160
; PRIORITY FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: AU PR1381
; PRIOR FILING DATE: 2000-11-10
; PRIOR APPLICATION NUMBER: US 60/245,596
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: PaesSTED for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1371

```

TYPE: DNA
 ORGANISM: Lawsonia intracellularis
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (1)... (1368)
 US-10-010-160-5

Query Match 5.7%; Score 269.2; DB 14; Length 1371;
 Best Local Similarity 52.1%; Pred. No. 3.3e-33;
 Matches 707; Conservative 0; Mismatches 633; Indels 18; Gaps 4;

2482 ACACAAATCCAGAAATTTTCAATTAACGGTGTATGCTGAACGTTCCATCCAAATTC 2541
 DB ATAAATACCAATAGCTTTTACCTTTCTATATAAATACCTTTCAAGTCCAAATAGCTTC 1300
 QY AGCAATTTGGTATACATTAACCTCCCAAAAGAGATACGAGATTTTAAATCTGTCTTC 2601
 DB AGCAATTTGGTATACATTAACCTCCCAAAAGAGATACGAGATTTTAAATCTGTCTTC 1240
 QY 1239 TGTATATCGGTAATATTTCTTGTATAGCATGTAATTTTCAAGTAAAAATTTGGTTC 1240
 DB 2602 AAATTTCTTACAGAGCTTTACGTAATGTACA--GAATAATCTTACACTTAAACATC 2658
 QY 1239 AAATTTCTTACAGAGCTTTACGTAATGTACA--GAATAATCTTACACTTAAACATC 1180
 DB 2659 ATTAATAGCAATAGATATATATCTAGTAAATCTTTGTCTGTATCATCTCTT 2718
 QY 1179 TTTCTTGGTAATATTTTAAATTTGCTTTTAACTTGAATCTGTCAAACTATCTTCT 1120
 DB 2719 AGGAAATTTCAATTAATTAATTCATTTCTTAACTTGAATCTGTCAAACTATCTTCT 2778
 QY 1119 CCCGTAATTAAGAAATTAACATTTCTGTCTAATAATTAAGTAATCTCTTACATCTCTGG 1060
 DB 2779 CCATTCATATGATGATCTGTATTAAGCTTCTTAACTTGAATCTGTCAAACTATCTTCT 2838
 QY 1059 CCATTCATATGATGATCTGTATTAAGCTTCTTAACTTGAATCTGTCAAACTATCTTCT 1000
 DB 2839 TATTTTCTTACAGATGATCTTATTAATTAATTAATTAATTAATTAATTAATTAATTA 2898
 QY 999 ATTAAGCTTTTACAGATGATCTTATTAATTAATTAATTAATTAATTAATTAATTAATTA 940
 DB 2899 TTTCTTCTTAAAGAGCTTCTTAAATTTGGAAGACATTTATTTATTAATTAATTAATTTCA 2958
 QY 939 TTACAGAAAGAGGGGAAATGCAATGAAACAACTCGTAAGGATTAATTAATTAATTAATTA 880
 DB 2959 AAACCTACAGCTTTTACCTTCAATATCTTGAAGAAAGACAAATTAATTTCTAC 3018
 QY 879 AAATTTCTTCAATGATGATCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 820
 DB 3019 ATCTATCTTAAAGAGATCTTAAATTTTCCCTAATTAATTTTCCCTCTGTAAATTAATCT 3078
 QY 819 ATCAATCTTAAATAGTTCTTAAAGATCTTAAATTTTCCAAATGTTCTTGAATAATAG 760
 DB 3079 GAGTATCTTAAATAGTTCTTAAAGATCTTAAATTTTCCAAATGTTCTTGAATAATAG 3138
 QY 759 CAAATTTTGT 700
 DB 3139 ATGATTTTGT 3198
 QY 699 TTTATGTGCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 645
 DB 3199 ACTTCTTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 3258
 QY 644 -CCTTTTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 586
 DB 3259 CATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3318
 QY 585 AACAGCTTAAATAGTTTGT 526
 DB 3319 TTTCTTCTTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 3378
 QY 535 CTCTTCTTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 466
 DB 3379 ATTAATCAATCTTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 3438

DB 465 TGATATAAATCTTTAAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 406
 QY 3439 A---AATGCTGATTTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 3495
 DB 405 AGTATATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 346
 QY 3496 CTTTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 3555
 DB 345 ATTAGCTGTAATTAAGATCTTTTCCAAAGAAAGAGCTTTTCAATTAATTAATTAATTAATTA 286
 QY 3556 CATATGCAAGATCTTACAGAGCTGTGCAATTAATTAATTAATTAATTAATTAATTAATTA 3615
 DB 285 TTGAGGATAGCTGTATACAGAGCTTTCAATTAATTAATTAATTAATTAATTAATTAATTA 226
 QY 3616 CAAATTAAGATACCTTTCTTAAAGCTTTTCCAGTACATTAATTAATTAATTAATTAATTA 3675
 DB 225 TAATTTCTGATGAAGAGCTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 170
 QY 3676 TCTTAACCATATATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3735
 DB 169 -GAAGCCAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 112
 QY 3736 GATTAACCATCAATCTGTAATTTAGTACATTAATTAATTAATTAATTAATTAATTAATTA 3795
 DB 111 TTTCTTGTGTAAGCTCTTTTAAATCTTCAATGAGGCTCATCTTCAAAATTTCTTCA 52
 QY 3796 TAGATTTCTGATATCAATCTTCAATCAATCAATCAATCAATCAATCAATCAATCAAT 3833
 DB 51 TGAAATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 14

RESULT 7

US-10-312-841-1/c
 ; Sequence 1, Application US/10312841
 ; Publication No. US2003018627A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Epigenomics AG
 ; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MK
 ; FILE REFERENCE: E01/1208/NO
 ; CURRENT APPLICATION NUMBER: US/10/312,841
 ; CURRENT FILING DATE: 2002-12-30
 ; NUMBER OF SEQ ID NOS: 2
 ; SEQ ID NO 1
 ; LENGTH: 3673778
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (3294164)
 ; US-10-312-841-1

Query Match 4.4%; Score 207.8; DB 14; Length 3673778;
 Best Local Similarity 45.4%; Pred. No. 3.9e-22;
 Matches 1233; Conservative 0; Mismatches 1447; Indels 37; Gaps 12;

QY 139 ACTGCTTCACTCCATGATATCTAGAAGCAACCAATTTCTCAATCTGCAATTTTA 198
 DB 171588 ATTCCTTAAACCACTCTTATTAACCAAAATCTATATCAATCAATTTTCTAAAAAACA 1715829
 QY 199 AAGTGAAGAAGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 258
 DB 1715528 AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1715469
 QY 259 AACTATCAAGCTCTTATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 318
 DB 1715468 AA-ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1715410
 QY 319 TCATTCATATTTTCCAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 378
 DB 1715409 CATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1715350

DB 1713206 CATATATAAATATATATATAAATATATATATATATATACACACACAAA 1713147
QY 2584 TTTTAAATACGTCTTTTCAAAATCTTCACAGCCTTACGTATGGTACAGAAATACCTT 2643
DB 1713146 TATAAATTAATACGTATATATATATATATATATATATATATATATATATAT 1713087
QY 2644 AGCACTTAAACATCATTAATAGGCAATTAAGATCTATCTACTGTAAATCTTTGC 2703
DB 1713086 AATATTAACAAA-----TAAATATATATATATATATATATATATATATATAT 1713031
QY 2704 TGTATCATCTCTTACAGAGATTGATTAATTAATTAATTAATTAATTAATTAAT 2763
DB 1713030 ATTAACAT 1712971
QY 2764 TCTTAAGTTACCTGGCATTGATGATGATGATGATGATGATGATGATGATGATG 2823
DB 1712970 ACTTATTAATCTAATCTAATCAATCAATCAATCAATCAATCAATCAATCAAT 1712911
QY 2824 ATGAGTACACAACTTA 2840
DB 1712910 AAAACAAAAAAACCA 1712894

RESULT 8

US-10-282-122A-16370
Sequence 16370, Application US/10282122A
GENERAL INFORMATION:

APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselebeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.034A

CURRENT APPLICATION NUMBER: US/10/282,122A

PRIOR FILING DATE: 2003-02-20

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/230,335

PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: 60/230,347

PRIOR FILING DATE: 2000-09-09

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/267,636

PRIOR FILING DATE: 2001-02-09

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 78614
SOFTWARE: Patent version 3.1
SEQ ID NO: 16370
LENGTH: 1017
TYPE: DNA
ORGANISM: Clostridium botulinum
US-10-282-122A-16370

Query Match 3.9%; Score 187.6; DB 12; Length 1017;
Best Local Similarity 52.1%; Pred. No. 2,6e-20;
Matches 499; Conservative 0; Mismatches 444; Indels 15; Gaps 3;

QY 635 AAATAGATATTTTAAAGTCAGAGATCATTTGGCACAGCAATATGCAATGCTGACGA 694
DB 31 AAGTTTGTATTTTAAAGACAGGATGATTTTGGACAGCTTTAGCTGTATGCTAGAAA 90
QY 695 CATGATATATCATTTAACTTATGGGAGCTGATCATGAAATATTTACATATTAACCT 754
DB 91 AAAGGATGAAAGTTAACTATGAGCAATGAAATATTAATATGATGAGATATTAATATC 150
QY 755 TACGAAAAAATTTAAATATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAAT 814
DB 151 AAAGAAAT 210
QY 815 AGCAATATATGAGAGATATATCTGACAAATCATGATATTTAACTATCTTAC 874
DB 211 AATGAAATGAAAAAGTATA--GAACATGATGAAATTTATGATTTTACAGTACCTTC 267
QY 875 CAACAATTAAGCACCATATGTAACAATAACAAACAAACAGCATATGTTAAATATCT 934
DB 268 CATGTATTAAGACAGGATATGCAAAAAGATC-----AACCTTTTAAAGAAATCA 321
QY 935 CCATATTAATTTGATAGTAAGGATATGAAATTTACATCATCAAAATTTCCAGTAAAT 994
DB 322 ATATATGATAGTATAGCTAAAGGATATGAGAGAGGTCAGGAAAAAGGTTATCAAGGTC 381
QY 995 GCAGAAAGAAATTTTCAATATATATCCATTTTATATCTCTGCTGCTGCTGCTGCTG 1054
DB 382 ATTAAGAAAGAACTGCTATATATCTTATATATATCTTATATATATCTTATATAT 441
QY 1055 GAAATGCAAAATCTTCTGATATATGATATGATGATGATGATGATGATGATGATG 1114
DB 442 GAAGTATGCTCAAGATATCTTACCAAGATGATGATGATGATGATGATGATGATGAT 501
QY 1115 GAATCATGATGAAACAAATAGTAAATGATGATGATGATGATGATGATGATGATGAT 1174
DB 502 GAAATGCTCAGGACCTTTTATGATATATATATATATATATATATATATATATAT 561
QY 1175 ATAGGTATACAGATGAGCTGCTCATTAAGAAATATGCAATTTGCAATGCTGGAATATC 1234
DB 562 GTAGAGGTAGAAATAGGAGCGCTGTAAATATATATATATATATATATATATATAT 621
QY 1235 GCTGAAAAAATTTTATGATATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1294
DB 622 GATGATATGATATGAGATATACAAAGAGCTTAAATGACAAAGATGATGATGATGAT 681
QY 1295 ATTAACACATATATATAGCAAAAAATCATTAATGATCTTCAATATATATATATAT 1354
DB 682 ATATATAGA-----ATAGTACAAAGCTTGAAGTAAACAGAAATCTTTTCAAGTCTT 735
QY 1355 TCATGCTTGGAGATCTATATATATATATATATATATATATATATATATATATAT 1414
DB 736 ACAGGTATGAGATTTAT 795
QY 1415 GCACTAGAAATAGGAAAAAGTGAATATATATATATATATATATATATATATATATAT 1474
DB 796 GGTATATCTATATGATTAAGGCTATGACACAGAAAGAAAGATTAAGAGATGATGAT 855
QY 1475 GTTGAAGAACAGATATGATTAACCACTGATATATTTGCAAAAAAATTTATATGATA 1534
DB 856 GTAGAGGGAATTAAGCTTGTAAAGCTTTTATATATATTAAGAAAGAAAGCTATGATGATA 915
QY 1535 CTACCAATTTGATATCTATATTTATATATATATATATATATATATATATATATAT 1592
DB 916 ATGCTTATTAACGAT 973

RESULT 9
US-10-282-122A-17353
Sequence 17353, Application US/10282122A
Publication No. US20040029129A1

GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See file wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: Patentin version 3.1
SEQ ID NO 17353
LENGTH: 1029
TYPE: DNA
ORGANISM: Clostridium difficile
US-10-282-122A-17353
Query Match 3.7%; Score 175.2; DB 12; Length 1029;
Best Local Similarity 51.6%; Pred. No. 2,4e-18;
Matches 483; Conservative 0; Mismatches 438; Indels 15; Gaps 3;

994 AGCAGAGAAATTTTACATATATATCAATTTTAACTCTCTGCTCAAGTTTGTCTAA 1053
369 CTGTGAAGAGAAATTCACAAAATCATATGTAATATATCTGAGGAGCACTCTATCTGA 428
1054 AGAAATTCAGACATCTCTCTGTATAGTACTGTGCTGATGAATATGAACCTTGG 1113
429 AGAAGTACAGAGGATATTTCTTACTACCGTAGAGTGTCTTCAAGATTAAATAG 488
1114 TGAATCATGTATGAACCAATTAATGATGCTTCAAAATATATATACCATTAAGATAT 1173
489 ACAGATGATCAAGATTTGTTATGAGCCCTTAACCTTAAGATTATACGAATTCAGACAT 548
1174 TATAGGTGACAGATTGAGCTGCATTTAAAGAACATATTCGAATTGCAATGGAATAT 1233
549 AGTTGAGTGAACCTTGAGAGGACATTTAAATATATATAGCATTTGGAGCTGGATAG 608
1234 CGCTGAAAAAATTTAGGTATATATATGCTGCTGCTGCTTATATAGTAAAGCATGATGA 1293
609 TGACGATTTAGGCTATATGATATATGCAAAAGCTGCTTATGACAAAGGATATGATGA 668
1294 AATTAAGACATATATATATAGCAAAATCATTCATCATATATATATATGCTTC 1353
669 AATGAGAGACTTGAATTC-----TATGGAGCTATATATCTTACCTTTGCAAGTCT 722
1354 ATCATGCTTGGAGATCTATATATTAATTAACATGTAACAGAACATTCACGAATATGCTTT 1413
723 TTGAGTATTTGAGATTTATATATGTTACTTGTCAAGTATGACATAGTAAGTAAGAGAGC 782
1414 TGACATGAAATATGAAAAAGGTAGAAATATTAATTAATATATGATACACAACTTAAGCT 1473
783 TGTATATCTATATGTAAGAGATGAGTTTATAGAGTACTTTAAAGAGTAAGATGCT 842
1474 TGTGAAGAACAGTACTGTAAACACTGATATCATTAAGCAAAAAAATTAATATAG 1533
843 TGTAAAGGATATAGCTACTAGTAAGTACATGATGACATGATGCTGAAAAATTAATATAG 902
1534 ACTACCATTTGCAATATCTATTTACAAATTTATTAACA 1569
903 TATGCTATATAGTATATATATCTGTAATATA 938
RESULT 10
US-10-282-122A-35500
Sequence 35500. Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09

PRIOR APPLICATION NUMBER: 60/242,578
 PRIOR FILING DATE: 2000-10-23
 PRIOR APPLICATION NUMBER: 60/253,625
 PRIOR FILING DATE: 2000-11-27
 PRIOR APPLICATION NUMBER: 60/257,931
 PRIOR FILING DATE: 2000-12-22
 PRIOR APPLICATION NUMBER: 60/267,636
 PRIOR FILING DATE: 2001-02-09
 PRIOR APPLICATION NUMBER: 60/269,308
 PRIOR FILING DATE: 2001-02-16
 Remaining Prior Application data removed - See File Wrapper or PAM.
 NUMBER OF SEQ ID NOS: 78614
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO: 35500
 LENGTH: 996
 TYPE: DNA
 ORGANISM: Staphylococcus haemolyticus
 US-10-282-122A-35500

Query Match 3.6%; Score 170.6; DB 12; Length 996;
 Best Local Similarity 49.6%; Pred. No. 1,3e-17;
 Matches 468; Conservative 0; Mismatches 469; Indels 6; Gaps 1;

635 AAAATCACTATTATAGTGCAGATCATTGGCAGACAAATGCAATGCACTGTCAGCA 634
 7 AAAATCTGTTTGGTATGGAAGTTTGGTACTGCTAGTAAATGATTAATGCTGAA 66
 695 CATGATATAGTAACTTATGCGGACGTGATCATGAATATATACATATTAACACT 754
 67 AATGACATACCTGTTTAAATGTTGGGAAAAAATGAGATGTTGAAAGATTTAAATGAT 126
 755 TACGAAAAATTTAAATATTTTACCCATATCTACAGACAAATATATGCAAC 814
 127 CACATCAAAATTAACGTTATTTAAAGATGCTTTAGATGATTAATTAAGCACT 186
 815 AGCAATATAGACGAATTTATCTGACAAATATCATGATTTATCTTAATCTTAC 874
 187 AGTATATTAAGAACCGCGAATTTTACTGATATATATTTAAAGGCTACCACTAAA 246
 875 CAACTATGACGACATATGTACCAAAATACAAACAAACGATATGTTAAAAATCT 934
 247 GCAATGCGGAAATTAAGTGAATGATTAATGATTTCTTAAGAAAACTTTAT 306
 935 CCAATATTAATTTAGTAAAGATATGCAATTTACATCACTCAATTTCCAGTGAATA 994
 307 CATGTGAAAAAGATATGAAATGATATCAATTCACAGCTTTTGAAGAAATGAT 366
 995 GCGAAGAAATTTTCAATATATCAATTTTATCTCTGCTGCAAGTTTGTAA 1054
 367 TCAATTTGAGAAATCACAATGATGATCGGGTTTATCAGGCTTACGCGCAGAA 426
 1055 GAAATTCAGAAACATCTCTGTTAGTATAGTACTGCTGCTGATTAATTAAGAACTGCT 1114
 427 GATGATGATTTAAACAAACATCACTACGCGCGCATCATTAAGATGAAAAATGAT 486
 1115 GAATCATGATGAAGAAACATTAATGATGCTTCAAAATATATATCAATCAAGATAT 1174
 487 AAACGATTCAGAAATTTATTTATGATGATATTTAGCTGATACATTAATGATTTA 546
 1175 ATAGTATGACATTTGAGCTGCACTTAAGAAACATTAATGCAATGATGGAATATC 1234
 547 GTCCGTTGTAATGAGGAGCATTTAAAAATTAATTTGCAAGTTGCAAGTTGATGTA 606
 1235 GCTGAAAAATTTAGTAAATGCTGTTGCTAATCTGTTATTAAGCAATGAATGAA 1294
 607 GACGACATGAGCTATGATGATATGCTTAAGCTCTTTAATGACACAGGTTTATGCTGA 666
 1295 ATTAAACATATATATGCAAAATATCATCAATATGATCTTCAATTAATTTATGCTCA 1354
 667 ATTAACGATTTAGTGAATTTAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720
 1355 TCATGCTTTGAGATCTATATTAATCAATGTAACAGAACTTACGCAATATGCTTTT 1414

Db 721 GGTGATTTGAGATCTAATTTGTTACTGTTACATCAACGATTCAGTAATTAACCTTTA 780
 Qy 1415 GGACTRGAATATGAAAAAGTGAATATTAATATATTAATGATCAACACTTAAGCTT 1474
 Db 781 GATTTAAATTTAGGACAAAGCCAAACAAATGATGAAGTTTAAATGAATGATATGCTA 840
 Qy 1475 GTTGAAGAACCACTGTTAAATCACTGATATCATTAAGCAAAAACTTAATGATGA 1534
 Db 841 GTAGAAGGATTTATTAATTAATCTGATATCAATTTAGTAAATCAACAAATGCTAT 900
 Qy 1535 CTACCAATTTGCAATCTATTTTACAAATTTATTAATGATGAATA 1577
 Db 901 ATGCCCTATTAATCAATGCACTTTTACAGGTTTATTCGAAGATA 943

RESULT 11
 US-09-815-242-8974
 Sequence 8974, Application US/09815242
 Patent No. US2002061569A1

GENERAL INFORMATION:
 APPLICANT: Haselbeck, Robert
 APPLICANT: Ohlsen, Karl L.
 APPLICANT: Zyskind, Judith W.
 APPLICANT: Wall, Daniel
 APPLICANT: Twick, John D.
 APPLICANT: Carr, Grant U.
 APPLICANT: Yamamoto, Robert T.
 APPLICANT: Xu, H. Howard
 TITLE OF INVENTION: Identification of Essential Genes in
 FILE REFERENCE: ELITRA 011A
 CURRENT APPLICATION NUMBER: US/09/815,242
 CURRENT FILING DATE: 2001-03-21
 PRIOR APPLICATION NUMBER: 60/151,078
 PRIOR FILING DATE: 2000-03-21
 PRIOR APPLICATION NUMBER: 60/206,848
 PRIOR FILING DATE: 2000-05-23
 PRIOR APPLICATION NUMBER: 60/207,727
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: 60/242,578
 PRIOR FILING DATE: 2000-10-23
 PRIOR APPLICATION NUMBER: 60/253,625
 PRIOR FILING DATE: 2000-11-27
 PRIOR APPLICATION NUMBER: 60/257,931
 PRIOR FILING DATE: 2000-12-22
 PRIOR APPLICATION NUMBER: 60/269,308
 PRIOR FILING DATE: 2001-02-16
 NUMBER OF SEQ ID NOS: 14110
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO: 8974
 LENGTH: 999
 TYPE: DNA
 ORGANISM: Staphylococcus aureus

FEATURE:
 NAME/KEY: CDS
 LOCATION: (1)...(999)
 US-09-815-242-8974

Query Match 3.5%; Score 168.6; DB 9; Length 999;
 Best Local Similarity 51.4%; Pred. No. 2.6e-17;
 Matches 505; Conservative 0; Mismatches 454; Indels 24; Gaps 4;

635 AAAATCAGTATTTTATGAGTCAAGATCAATTTGGCAGACAAATGCAATTTGCACTGTCAGCA 694
 7 AAAATACCGTTTGGTATGGAAGTTTGGGACAGCCCTTGCATGTTCTTCCAGAA 66
 695 CATGATATCACTTAATTTAGGAGACGTGATCATGAAATATATATATATATATATATAT 754
 67 AATGACATGATGTTTATGATGTTGGGTTAAAAATCAAGATGCTTTGATGATTAATATCA 126
 755 TACGAAAAATTTAAATATTTTACCAATATCATCTACAGACAAATATATATGCAAC 814
 127 TGTATACAAATTAAGATTTAAATATGAGAAATTAAGATGATTAATCATCATCTACT 186

```

APPLICANT: Forsyth, R.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: EMTPA.034
CURRENT APPLICATION NUMBER: US/10/282,122A
PRIORITY FILING DATE: 2003-02-20
PRIORITY APPLICATION NUMBER: 60/191,078
PRIORITY FILING DATE: 2000-03-21
PRIORITY APPLICATION NUMBER: 60/206,848
PRIORITY FILING DATE: 2000-05-23
PRIORITY APPLICATION NUMBER: 60/207,727
PRIORITY FILING DATE: 2000-05-26
PRIORITY APPLICATION NUMBER: 60/230,335
PRIORITY FILING DATE: 2000-09-06
PRIORITY APPLICATION NUMBER: 60/230,347
PRIORITY FILING DATE: 2000-09-09
PRIORITY APPLICATION NUMBER: 60/242,578
PRIORITY FILING DATE: 2000-10-23
PRIORITY APPLICATION NUMBER: 60/253,625
PRIORITY FILING DATE: 2000-11-27
PRIORITY APPLICATION NUMBER: 60/257,931
PRIORITY FILING DATE: 2000-12-22
PRIORITY APPLICATION NUMBER: 60/267,636
PRIORITY FILING DATE: 2001-02-09
PRIORITY APPLICATION NUMBER: 60/269,308
PRIORITY FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 76614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 15395
LENGTH: 999
TYPE: DNA
ORGANISM: Clostridium acetobutylicum
US-10-282-122A-15395

Query Match          3.5%; Score 168.2; DB 12; Length 999;
Best Local Similarity 51.0%; Pred. No. 3.1e-17;
Matches 482; Conservative 0; Mismatches 448; Indels 15; Gaps 3;

QY 648 TAGGTGCAGATCATTTGGGCACAGCAATAGCAATGTCACGTGACGCATGCTATATCAG 707
DB 20 TAGAGAGAGAGAGCTTTGGTACAGCACTTCGCAATATGCTTGCTAAAAAGGTATATG 79
QY 708 TTAACATATGGGACGTGATCATAGAAATATTAACATATTAACCACTTACCGAAAAAT 767
DB 80 TTGTGATTTGGGATAGAAATTAAGAAATTTCTTGAGATATTAATATCAATTAAGAACTATA 139
QY 768 TAAATATTTTACCACATATCATCTTACCAAGACAAATATATGCAACGACCAATTAACG 827
DB 140 CAAATATATCTTCCAAACAACTAATATACCTGCTGCGTTAAAGCAGTAGATATATGGA 199
QY 828 AAGATATATCTGACAAACAAATACATGTATATCTTAATATCTTACCAACAAATTAACGA 887
DB 200 AGG---CAGCTAAGGAAGAAATATATGTCTTCTGCTGCCCATCTTTTGCTAAGGG 256
QY 888 CCAATGTACCAAAATACACAAACAAAGACATATGTGTAAAAAATCTCCATATTAATTT 947
DB 257 AAGATATCAGAAATATTAAGGGCTT-----TTTAAGAGAAACCAATATATTTTAAGTA 310
QY 948 GTAGTAAAGGTATCGAATATTAATCACTCAAAATTTCCAGTGAATAGCAGAGAAAAATT 1007
DB 311 TTGCAAAAGGCATGGAAGAGAGACCAAGAAAGGCTGTCTGAAGTATGTTAAAGAGAAC 370
QY 1008 TAGCAATATATCCAAATTTTAAATCTCTGTGTCACAAATTTTGGCTAAAGAAATGCAAGAC 1067
DB 371 TTTATTAATAATCTGTGTGTAGTACTTTTCTGTGCTTAATCATGACGAAGAAATTCGAATG 430
QY 1068 ATCTTCTTGTAGTATAGTACTTGTCTGTGATATAAAGAACTTGTGATCATTTGATG 1127
DB 431 ATATTTCCAACTACAGTATGTGTTACTTCCACAGATATGAATATGACGAAGAAATTTCAAG 490
QY 1128 AAACAATATAGTAAGATGTTCTAAAAATATATATCCATCAAGATATATATAGGTATACGA 1187

```


QY 1302 CACTATATATAGCAAAAAATCTATCATATGATCTTCATCATTAATGGTCCATCATGTC 1361
DB 725 GTCTGGGTGTCGCGATGGGT-----GCCAATCCCTGACATTTTACGTTTGAGTGCTG 778
QY 1362 TTGGAGATCTATATTAATACATGATACAGCAAGACATTCAGCAATATGGCTTTGGACTAG 1421
DB 779 TGGGGGACTGATCGTACTTGTACAGAGTTCATTCGCAATGGGCTGCCGTA 838
QY 1422 AATAGGAAAGAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1481
DB 839 TGTAGGAGAGGAGCAATAGCTGAGAGAGTCTGAGAGAGAGAGAGAGAGAGAGAGAG 898
QY 1482 GAACAGTACTGTAAAGCAAGATATCATGATATCATGATATCATGATATCATGATATCAT 1541
DB 899 GTGTGCAACGACAAAG 958
QY 1542 TTGCAATATCTATTTTCAATTTTATTAATTAATTAATTAATTAATTAATTAATTAAT 1601
DB 959 TTAATCAAACTATATATGTTTGTATTAATGAGAGAGATCAAAAAAGAGAGAGAGAG 1018
QY 1602 ACATATAT 1610
DB 1019 AATTCATGT 1027

RESULT 16
US-10-311-455-1369/c
; Sequence 1369, Application US/10311455
; Publication No. US20030143606A1
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIERENBERG, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining the Methylation Status of Cytosine Residues
FILE REFERENCE: 5013.1014
CURRENT APPLICATION NUMBER: US/10/311,455
CURRENT FILING DATE: 2002-12-16
PRIOR APPLICATION NUMBER: PCT/EP01/07537
PRIOR FILING DATE: 2001-07-02
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 2424
SEQ ID NO 1369
LENGTH: 5641
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1369

Query Match 3.3%; Score 155; DB 14; Length 5641;
Best Local Similarity 43.9%; Pred. No. 7,4e-15;
Matches 1027; Conservative 0; Mismatches 1280; Indels 34; Gaps 7;

QY 176 ATTTTCTCAATCTCCATTTTAAATGATGACAGAGATTTTAAATGATGACAGAGATTTTAA 235
DB 4946 ATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4887
QY 236 TACTCTTATGAG 295
DB 4886 CACCCCAATTTTAAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4830
QY 296 TAAAGAAATACCT 355
DB 4829 AGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4770
QY 356 GTTCTTATGAG 415
DB 4769 TAAATCTATTTTCAATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCT 4710

QY 416 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 471
DB 4709 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4650
QY 472 ACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 531
DB 4649 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4590
QY 532 TGAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 591
DB 4589 ATATCTATCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4530
QY 592 ATATTTTGGAG 651
DB 4529 ATATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4470
QY 652 TGCAGATCAATTTGGAG 711
DB 4469 TAAATCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4410
QY 712 CTATGAG 771
DB 4409 ATAT-----AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4357
QY 772 ATATTTACCCATATCATCTATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 831
DB 4356 AATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4297
QY 832 ATATCTGAG 891
DB 4296 ATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4237
QY 892 ATGTAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 951
DB 4236 ATTAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4177
QY 952 TAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1011
DB 4176 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4117
QY 1012 ATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1071
DB 4116 CTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4057
QY 1072 TCCTTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1131
DB 4056 AATCTATCTAATTTTCAATCTAATTTTCAATCTAATTTTCAATCTAATTTTCAATCTAATTTT 3997
QY 1132 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1191
DB 3996 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3937
QY 1192 AGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1251
DB 3936 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3877
QY 1252 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1311
DB 3876 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3817
QY 1312 AGCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1371
DB 3816 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3767
QY 1372 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1431
DB 3766 ATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3707
QY 1432 AGGTAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1491
DB 3706 AATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3651

QY 1442 ATAAATACATTAATGATGACCACTAAAGCTTTGTAAGGACGATCTGTAACCA 1501
DB 318 ATGGATCAAGCATTTATCTGAAATGATATGTTGTAAGGATTTATTAACAACCA 259
QY 1502 CTGATATCATTTGCAAAAAAATCTATATGAGAACCAATTTGATATCTATTAACAT 1561
DB 258 GTTATCATTTAGCTTAAGAAAAAATGTGATATGCCAATTCACAAATGATATATGA 199
QY 1562 TTATTCATGAGAAATATATCTAGATTAAGCCATATCAACATTTATCTAGCTATC 1621
DB 198 GATATTTTGAATAATCTCAGTAAAGATGCGTAAAGATTTATGAGCGCATAAA 139
QY 1622 ATCTTTGATATTTTCAACCAATGATTAATTAAGGTCGAAAT 1666
DB 138 AAATCTGAATTAATGTGTATTTATTCACATATACGGCTAATAT 94

RESULT 18
US-10-329-624-513/c
Sequence 513, Application US/10329624
Publication No. US20040043037A1
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
Gill H. Choi
Patrick S. Dillon
Craig A. Rosen
Steven C. Barash
Michael R. Fannon
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5256
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/329,624
FILING DATE: 27-Dec-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/956,171
FILING DATE: October 20, 1997
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 513:
SEQUENCE CHARACTERISTICS:
LENGTH: 918 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 513:
US-10-329-624-513

Query Match 3.2%; Score 154.2; DB 12; Length 918;
Best Local Similarity 53.6%; Pred. No. 4.9e-15;
Matches 346; Conservative 0; Mismatches 293; Indels 6; Gaps 1;

QY 1022 ATTTTATACCTCTGCTCCAGTTTGGCTAAAGAAATTGCAACATCTCTCTAGT 1081
DB 732 ATGGCGGTGTTGACAGGCCAGTCTATGCGAAGAGTTGTGTCAAGAACCACTACA 673
QY 1082 ATAGTACTGCTGCTGATATATAAGAACTTGGATCATTTATAGAAACAAATTAAT 1141
DB 672 GTTGTGCTTCATCAACAAAGATTAAGTAAATTAACGAAATTTATTTATGAAAT 613
QY 1142 GATGTTCTAAAAATTAATATATCAATCAAGATTTATAGGTGACAGATTTGAGCTGATTA 1201
DB 612 GATTTATTTGGTGTGACACGAATGATGATCTGATTTGGTGTGAACCTTGCTGATTTG 553
QY 1202 AAGAACATAATTCGAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1261
DB 552 AAAAAATATCTGCGATAGCAAGTGTATCTGATGATGATGATGATGATGATGATGAT 493
QY 1262 GTTGTCTGCTGTTTAACTAAAGGATGAATGAATTAACACTATATATAGCAAAAAT 1321
DB 492 AAAGCTGCATTAATGATCTGCTGCTTACGGAATTAATGATGA-----TTAGGTGAAAAG 439
QY 1322 CATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1381
DB 438 TTAGTGCGCATCTTATGACATTTCTAGGTTTATGATGATGATGATGATGATGATGAT 379
QY 1382 TGTACACAGACATTTACGCAATATGCTTTTGAATGAAATAGCAAAAGGTAGAAAT 1441
DB 378 TGCACATCAACACATCTCGGAATTTACATTAAGATATATACTTGACAAAGGTAGTCA 319
QY 1442 ATAAATACATTAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1501
DB 318 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 259
QY 1502 CTGATATCATTTAGCAAAAAAATCTTAATGATGATGATGATGATGATGATGATGAT 1561
DB 258 GTTATCATTTAGCTTAAGAAAAAATGTGATATGCCAATTAACAATCATATATATGA 199
QY 1562 TTATTCATGAGAAATATATCTAGATTAAGCCATATCAACATTTATCTAGCTATC 1621
DB 198 GATATTTTGAATAATCTCAGTAAAGATGCGTAAAGATTTATGAGCGCATAAA 139
QY 1622 ATCTTTGATATTTTCAACCAATGATTAATTAAGCGTCGAAAT 1666
DB 138 AAATCTGAATTAATGTGTATTTATTCACATATACGGCTAATAT 94

RESULT 19
US-10-311-455-2128/c
Sequence 2128, Application US/10311455
Publication No. US20030143606A1
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PREPNBROCK, Christian
TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Deter
FILE REFERENCE: 5013.1014
CURRENT APPLICATION NUMBER: US/10/311,455
FILING DATE: 2002-12-16
PRIOR APPLICATION NUMBER: PCT/EP01/07537
FILING DATE: 2001-07-02
PRIOR APPLICATION NUMBER: DE 10032529.7
FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 2424
SEQ ID NO 2128
LENGTH: 15548
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-2128

Query Match	3.2%;	Score 153;	DB 14;	Length 15548;
Best Local Similarity	44.7%;	Pred. No. 2.3e-14;		
Matches 981;	Conservative	0;	Mismatches 1185;	Indels 27;
				Gaps 9;

QY	237	ACCTTATGCAAGAACTCTTAAACTTATCAAGTCTCTTACCTAAGAAACAACGTGTT	296
Db	13534	ACACATCGATTAATAATATATATATTAATAACACATACGTATTAATAATATACGGATATATATA	13475
QY	297	CAAGAATACCTCTCTCTTGTTCCATCCATATTTTCCATTAATTTCACATATCATAGG	356
Db	13474	CATATATAAACACATACGTATTAATAATATACATATAAAATATACATATATATA-AAACATATA	13416
QY	357	TTCTTTATTCGACATACATATTAATAATTCATCCAAAACAAATTAATAGTACTATCTATTA	416
Db	13415	ATATATAATATATATACGTATTAATAATATACACGGCATTAATAATATATACGTATTAATAATATAT	13356
QY	417	ATATAAAAAATCCCAATAGTCTTTAGCAGAAGCTACAAATATATATAGACACTTAACCTTA	476
Db	13355	ATATATAATATACACGTATAAAATATATATATATATAAAATTTCAAGTATATAAATATATAT	13296
QY	477	TTATATATATTCCAATTTTATTAATCTACTACATCATATATTTAGCATATATACATCTGCA	536
Db	13295	CTATAAATATATATATATAAATATATATCTATAAATATATATATAAATATATATCTGTAATA	13236
QY	537	TGATTTAACATTAAGTAACACATTTCTTACACTGTAGATATATACATAAATAATATATAT	596
Db	13235	TATATATATATATA-AAATATATATCTATAAATATATATATATAAATATATATCTTAATAAT	13177
QY	597	TTTTGACGTGTTCTTATATATAATTAAGTATAAATAATGAAAAATCAGTATTTTAGTGCA	656
Db	13176	ATATATATAATATATATATCTATAAATATATATATATATATAAATATATATCTTAATAAT	13117
QY	657	GATCATTTGGCACGCAATAGCAATTGCACCTGTCAGCACATGATATATCAGTTAACTTAT	716
Db	13116	ATATAAATATATATATAAATAATATATATAAATAATATATATAAATAATATATATATAA	13057
QY	717	GGGACGTGATTCATAGAAATATTTACACATATAAACACTTACCGAAAAATTTTAAATAT	776
Db	13056	TATATATATACATATATACATATATATAAATATATATATATATATATATATATATATATATAC	12997
QY	777	TACCCACATATCATCTTACCGACACACATATATGCCACAGCAATATAGCGAAATTTAT	836
Db	12996	TATATACATATATATATATAAATAATATATAAATAATATATATAAATATATATATATAT	12937
QY	837	CTGCACAATATACATGATATATCTTAACTATCTCCACACACAATATAGCCACCATATGTA	896
Db	12936	ATATATTCGAAATATATA-TATACGAAATATATATATATATATATATATATATATATATA	12878
QY	897	CACAAATACACACAAACAGCATATGTTAAAAATCTCCATATTTAATTGTAGTAAG	956
Db	12877	TACAAAT	12818
QY	957	GTATCGAAATTAACACTACCTCAAAATTTCCGTAATATAGCAGAAATTTTACATATA	1016
Db	12817	ATA	12758
QY	1017	ATCCAAATTTTATATCTCTGTGTCAGATTTTCTTAAAGAAATTCAGAAATCTTCTCT	1076
Db	12757	TATATATATTCGAAAT	12698
QY	1077	GTAGTATAGTAACTGCTGGTATATATATAAAGCTGTGTAATCATATGATAGAAACATATA	1136
Db	12697	TAT	12638
QY	1137	GTAAATGATGTTCTTAAAAATA-----ATATACATCAAGATATATATAGTGTACAGAT	1188
Db	12637	ATATATTCGAAAT	12578
QY	1189	TGGACGTCAATTAAGAAACAATAATTCGAATTCGATGCGAATAATGCGTGAATAAATTT	1248
Db	12577	ATATAAATATATACATATAAATA	12518
QY	1249	AGGTATATATGCTGTGCTACTGTATTAACATAAAGCATGAATGAATTAATAAACAATA	1308

[illegible]

Db 956 AACAGCAGCAAAAGAAATCATGT 979

RESULT 23
US-09-815-242-3869
Sequence 3869, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Heselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT FILING DATE: 2001-03-21
PRIOR FILING DATE: 2001-03-21
PRIOR FILING DATE: 2000-03-21
PRIOR FILING DATE: 2000-03-21
PRIOR FILING DATE: 2000-05-23
PRIOR FILING DATE: 2000-05-23
PRIOR FILING DATE: 2000-05-26
PRIOR FILING DATE: 2000-05-26
PRIOR FILING DATE: 2000-10-23
PRIOR FILING DATE: 2000-10-23
PRIOR FILING DATE: 2000-11-27
PRIOR FILING DATE: 2000-12-22
PRIOR FILING DATE: 2000-12-22
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3869
LENGTH: 1026
TYPE: DNA
ORGANISM: Enterococcus faecalis
US-09-815-242-3869

Query Match 3.1%; Score 147.2; DB 9; Length 1026;
Best Local Similarity 49.1%; Pred. No. 6,6e-14;
Matches 480; Conservative 0; Mismatches 488; Indels 10; Gaps 3;

QY 635 AAAATCAGTATTTAGTGCGAGATCATTTGGCAGCAGCAATGCAATTCAGCTGTGACA 694
DB 4 AAAGTCGCTGTTTGGTTCAGGTTCTGGGGGACAGCTTTAGCTCAAGATTAGCGGAA 63
QY 695 CATGTTATATCATTAATCTATGAGGAGCTGATCATATAAATTTACATATTAACACT 754
DB 64 AATGACACGAGGCTGTATTTGGGAAATAAACAAGCAAAATCGATGAATATAACA 123
QY 755 TACCGAAAAATTTAAATTTATTTACCATATCATCTACCAAGCAACATATATGCAAC 814
DB 124 AAACACACCAAAACATTTATTTACAGATTTAATCTTACTTCAATTCAGGCAACC 183
QY 815 AGCAATATAGAGAGATTTATCTGACAAACATAC--ATGATTTATCTTAATCTTCTCTA 872
DB 184 ACTGATTTAGGCACTGCTTTAGTAGATGTGATGGGAGTCTGTTGGTACCAACTAAA 243
QY 873 CACAACAATTAGCAGCATATGTACACAAATACAAACAACAACAGCATATGTATAATA 932
DB 244 GCCATTCGTGATGACACAAAGAGTGGCCCAACCTTTAAAAAACAACAATATTTAT 303
QY 933 CTCGAATTTAATTTGTATGTAAGGTATCGAAATTTACATCACTCAATTTCCCGTAAA 992
DB 304 CATGCAAGTAAAGGCTTGAAACAAGGTACACATTAAGGCAATTTGAGAAGTGAATGCGGAA 363
QY 993 TAGCAGAAGAAATTTACATATATCAATTTTATCTGTGTCGAAGTTTGCTA 1052

Db 364 GAGATTCAGACAA--AAAACGCAAGGATTTGCTTTTATCTGTCCAGCAAGCAG 421

QY 1053 AAGAAATTGCAAGACATCTTCTGTAGTATAGTACTGCTGTGATATAAAGACTTG 1112
DB 422 AAGAGTCGCTGTTCATGACATTTACACACATTTACAGCAAGCAAGCAAAATTTAGCAGAG 481
QY 1113 GTGAATCATTTGATGAAACAATTAAGTATGATGCTTTTAAATAATATATCCATCAAGATA 1172
DB 482 CTGTGTATGTCAGAGAGTGTGTTATGATGATTAATCTTTGAAATTTATACAAATGATGATG 541
QY 1173 TTATAGGTGACAGATTTGAGCTGATTAAGAAACATAATTCATTTGATGATGATGATTA 1232
DB 542 TGATTTGTGTAAGAAACAGGGGCGCTTTAAAAATTTATTTGCTTTGGGCGCCGAGCA 601
QY 1233 TCGCTGAAAAAATTTAAGTATATATGCTGTTGCTACTGTTATTAATTAAGGCAATG 1292
DB 602 TTGATGCTTTAGGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 661
QY 1293 AAATTAACACATATATATAGCAAAATATCATTCATTAATGATGATGATGATGATGATG 1352
DB 662 AAATTTAGCTTTAGGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 715
QY 1353 CATCATGCTTTGAGATCTTAATATTAATGATGATGATGATGATGATGATGATGATGATG 1412
DB 716 TAAAGTGTGTTGGGATTTAATTTATCTTTGATGATGATGATGATGATGATGATGATG 775
QY 1413 TTGAGCTAGAAATAGGAAAGGATGAAATATTAATCATTAATGATGATGATGATGATGATG 1472
DB 776 CGGTATTTATTTAGGAAAGGCAATTAATGATGATGATGATGATGATGATGATGATGATG 835
QY 1473 TTGTTAGAGAACAGTACTGTAAACCACTGATCATTAAGCAAAAAATTAATGATG 1532
DB 836 TTGTGAGAGGTGTTTCAACAACCTAAAGCAGCTGATGATGATGATGATGATGATGATG 895
QY 1533 AACTACCAATTTGATATCTATTTTAAATTTTAAATGATGATGATGATGATGATGATG 1592
DB 896 AAATGCAATCAGCAAGCAATCTTACATGCTTTGATGATGATGATGATGATGATGATG 955
QY 1593 CCATATCAAAACATTTAT 1610
DB 956 CAGCAAAAGAAATCATGT 973

RESULT 24
US-10-311-455-836/c
Sequence 836, Application US/10311455
Publication No. US20030143606A1
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPENBROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Detect
FILE REFERENCE: 5013.1014
CURRENT FILING DATE: 2002-12-16
PRIOR FILING DATE: 2002-12-16
PRIOR FILING DATE: 2001-07-02
PRIOR FILING DATE: 2001-07-02
PRIOR FILING DATE: 2000-06-30
PRIOR FILING DATE: 2000-06-30
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 2424
SEQ ID NO 836
LENGTH: 6132
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-836

Query Match 3.1%; Score 147.2; DB 14; Length 6132;

Best Local Similarity 44.2%; Pred. No. 1.3e-13;
Matches 1069; Conservative 0; Mismatches 118; Indels 31; Gaps 10;

QY 281 AGAAGACAGCTGATTCAGAAAGATACCTCTCTCTGTCATCCATTTTCCATTTAT 340
Db 5037 AACCAAAAACCTTTATTTCTTCAACCTCTTAAATTTTCTCCCAAAATCCCTAAAAAT 4978
QY 341 TTACAAATGATAGAGTCTTTATTCGACATACATATTAATCAATCAAAACAAATTA 400
Db 4977 ATAAAAATATATATA---TTTATTAACACTATATTAATTAATTAATTAATTA 4921
QY 401 TAGACATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 460
Db 4920 AAAACATTAACCACTTAAAAAAATTAATAAACTAATAAAAAAACTATCTACAT 4861
QY 461 TGAACACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 520
Db 4860 TTATACAC--ACAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4802
QY 521 AATATCATCTGATGATGATTAACATATGATCACTTTCTTACCTGATTAATTA 580
Db 4801 ACTACCAAAACAACTCAAAAAAAATTAATAATCAACCAATTAATTAATTAATTA 4742
QY 581 ATAAATATATATATATTTTGAATGCTTCTGATTAATTAATTAATTAATTAATTA 640
Db 4741 CTATTCATATTTAC--TAAATTCCTTAACAAACAAATTAATTAATTAATTAATTA 4684
QY 641 AGATTTTAGGTGAGATCATTTGGACAGCAATGCAATGCACTGTCAGCATGCT 700
Db 4683 AAAATATCCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4624
QY 701 ATATCACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 760
Db 4623 ATCAAAAAAAATCCCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4564
QY 761 AAAATTTAAATATTTTACCAATATCATCTACAGACAACTATATGCAACCAAGCA 820
Db 4563 CAATTTTCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4504
QY 821 ATGAAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 880
Db 4503 AAAAAAAACCTAAA---AAACATTTCTTTATTAATTAATTAATTAATTAATTA 4448
QY 881 TTACGACCATATGATCAAAATTAACACAAACAGCATATGTAATTAATTAATTA 940
Db 4447 ATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4388
QY 941 TTATTTGATTAAGATGATGATTAATTAATTAATTAATTAATTAATTAATTA 1000
Db 4387 CTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4328
QY 1001 GAAATTTTACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1060
Db 4327 CAATATCTCCCAAAACATCTAATTAATTAATTAATTAATTAATTAATTAATTA 4278
QY 1061 GCGAAGATCTTCTGATTAATTAATTAATTAATTAATTAATTAATTAATTA 1120
Db 4277 TCAATTAATTTCTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4218
QY 1121 TTGATTAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1180
Db 4217 AAATCCAAAAAAACAAACATATTAATTAATTAATTAATTAATTAATTAATTA 4158
QY 1181 GTACAGTGGAGTGTATTAAGAAATTAATTAATTAATTAATTAATTAATTA 1239
Db 4157 TCTCAATTTCAAACTTACAACTACAAATTAATTAATTAATTAATTAATTA 4098
QY 1240 AAAAAATTTAGTAATATGCTGTGCTACTGTTATTAATTAATTAATTAATTA 1299
Db 4097 AAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4038
QY 1300 AACCTATATTAAGCAAAAAATCATTCATTAATTTCTTAAATTTGCTCATG 1359

Db 4037 TTTAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3978
QY 1360 TCTTGGAGATCATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1419
Db 3977 TCTTCAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3918
QY 1420 AGAAATAGAAAAAGTGAATATTA--TACATTAATTAATTAATTAATTAATTA 1477
Db 3917 CTACCTATATACATCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3858
QY 1478 GAAGAACAGCTACTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1537
Db 3857 AABCCATTAATCTTTAAAAAAACAAAGTAATCTTCAATTTTAAATTTATTA 3798
QY 1538 CCAATTTGATATCTATTTAATTAATTAATTAATTAATTAATTAATTAATTA 1597
Db 3797 AATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3738
QY 1598 TCAACATATTAATCTTATGCTATGATCTTGAATTAATTAATTAATTAATTA 1657
Db 3737 ACATTAATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3678
QY 1658 GTCAAAAAATATAGATCAACGTTCAACCATATATTAATTAATTAATTAATTA 1717
Db 3677 TACAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3618
QY 1718 TAACTCACTAGTATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 1777
Db 3617 ATATTAACAACTCTTCAACCTCAACAAATTAATTAATTAATTAATTAATTA 3558
QY 1778 CTACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1837
Db 3557 CAAAAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3498
QY 1838 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1897
Db 3497 AATATATCTCAACCTCTTCTCAATTAATTAATTAATTAATTAATTAATTA 3438
QY 1898 CACTACATCATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1957
Db 3437 CTTTACACCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3378
QY 1958 TACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2017
Db 3377 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3318
QY 2018 ATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2075
Db 3317 AAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3258
QY 2076 TTATGCCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2135
Db 3257 CCAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3198
QY 2136 CTGACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2195
Db 3197 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3138
QY 2196 AACCACTACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2255
Db 3137 ACCCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3078
QY 2256 GCATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2315
Db 3077 TTATTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3018
QY 2316 ACTTCAACAACTCAAAAGTTCTATTAATTAATTAATTAATTAATTAATTA 2375
Db 3017 AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2295
QY 2376 CAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2435
Db 2295 TATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2899

Best Local Similarity	49.8%;	Pred. No. 2.1e-12;							
Matches	484;	Conservative	0;	Mismatches	473;	Indels	14;	Gaps	4

QY	613	TATTTAAATTTAGTAAAAAATGAAAAATGAAATGATTTTAAAGTGCAGAGATCATTTGGCAGCG	672
Db	2065487	GAAAAATTTATATGACACAGAAAAAAGTGCATTTCTTTGGCCCTGGAAATTTGGGAAACAGG	2065428
QY	673	AATAGCAATTTGCACCTGTGACGACACATGGTATATCACTTAACCTTATTTGGGAGCTGATCATG	732
Db	2065427	ACTGCACCTTGTCTTAAGCTGATTAATAATCATTAACCACTTAATTTGGGAAACTTAGTTAA	2065368
QY	733	AAATTTACACATATATAACACATTACCGGAAAAATTTAAATTTAAATTTCCCATATTCATCT	792
Db	2065367	AATTTGATGAAATTTAACGAAATGCGACACGAAATGTGACTATTTGCCAGATATATT	2065308
QY	793	ACCGACACAATATTTGCAACCGAATATTTAGAGAAATATTTATCTGCAACAATACATG	852
Db	2065307	ACCAACTGAGGTAAGGAAAGGACACTTATCACTTGATGAACCTTAAGATGTGCTGAAT--	2065251
QY	853	TATTAATCTTAATCTATTTCTTACACACAATTAACGACCATATGTACACAAATACACACAA	912
Db	2065250	TGTGCTGATTTGCTATTTCCACAAATGCAATGCGTGTGTTTGTAAAGCAGTAAATGAAGC	2065191
QY	913	ACAGATATGTTGTAATAATTAATCTCCCAATTTATTTATTTGTAAGTAAGTATCGAAATTAATC	972
Db	2065190	GCTGAA-AGAACCAATATTTTATGTCATGTTAGTAAGGATGAACACGAAACCAATTC	2065132
QY	973	ACTCAATTTTCCAG---TGAATAGCAGAAAGAAATTTTCAATATATATCAATTTT	1028
Db	2065131	TTAGATATCTAGAAAGTAATTTAGAGAGAGATGATCAACAAACGTAACCTCTGTG	2065072
QY	1029	TACTCTCGTGCAGGATTTTGTCTAAGAAATTTGACGAACATCTTCTGTGATATAGAC	1088
Db	2065071	TTCTTTTGGGGCTTGTACCGCTGAGAGATGAGCTCTTCTGCAATCAACAAACGCTTTGTG	2065012
QY	1089	TTGCTGTGATTAATPAAGAACTGTGTGATCATTTGATGAACAATAGATATGTTTC	1148
Db	2065011	CTAGCTGTAAGATTTTGAAACAGACAGAAATGCGAAATGCTTTTATTAATTAATTT	2064952
QY	1149	TAAATAATATATACCATCAAGATTTATAGGTGTACAGATTTGAGCTGCATTAAGACA	1208
Db	2064951	TACGATTTTATACGAATATATATGTCATCGGTCTGAAATTTGTGTGACGCTTTAAAAATA	2064892
QY	1209	TAAATTCATTTGCATGTGCAATTAATCGCTGAAAAAATTTAGATATATGCTGTGCTA	1268
Db	2064891	TTATTCGCACTTTGGCGCAGAGAACTCTGTATGCGCTTGATACGGGATATATCTTAAACGGG	2064832
QY	1269	CTGTTATATACATAAAGGCGATGATGAATGAATTAATAACATATATATAGCAAAAAATCATTCAA	1328
Db	2064831	CACATATATACAGCGGGAATGCTGAAATTAATCTGCTCGAGATGTGTGTGGGCTCTA---	2064775
QY	1329	TAGATCTTCATACATTAATTTGTGTCATATGTCTTTGAGATCTAATATTAACATGTACAA	1388
Db	2064774	--ATCCACAACCTTTTACGGGCTTAATGTGATATTTGACCTTAATCTTATCTGTACAA	2064718
QY	1389	CAGAACATCTACGCAATATGCTTTTGGACATAGAAATAGAAAAAGTATGAATTTAATA	1448
Db	2064717	GTTGTCATCTACGTAATTTGGGTGTGTGAAACATGTCTGGTAAAGGAGAAATTTAGATG	2064658
QY	1449	CATTTAATGATCAACAACCTTAAGCTTTGTGAAGAACAGATATCTGTAAACCACTGATAT	1508
Db	2064657	AAGTTTATGAAAAAATGGGATATGGTGTGAAGGTTGTCGTAACGCCAAGCAACCTTCATG	2064598
QY	1509	CATTAGCAAAAAAATTTAATGTAGAACCTACCAATTTGTGATATCTAATTTACAAATTTATAC	1568
Db	2064597	GTTGGGCGAAAAAACTAGATATTTGTATGCAATTCAGAAATCGATTTATCGCAGATTTTAT	2064538
QY	1569	ATGAGAAATATA 1579	
Db	2064537	TCGAGAAATPAA 2064527	

US-10-240-453-206/c

```

! Sequence 206, Application US//10240453
! Publication No.: US20030148326v1
! GENERAL INFORMATION:
! APPLICANT: OLER, Alexander
! APPLICANT: PIERENBERG, Christian
! APPLICANT: BERLIN, Kurt
! TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA
! TITLE OF INVENTION: Transcription
! TITLE OF INVENTION: by Means of Assessing the Methylation Status of Genes Associated
! FILE REFERENCE: 5013.1009
! CURRENT APPLICATION NUMBER: US/10/240,453
! PRIOR FILING DATE: 2002-10-02
! PRIOR APPLICATION NUMBER: PCT/EP01/03973
! PRIOR FILING DATE: 2001-04-06
! PRIOR APPLICATION NUMBER: DE 10019058.8
! PRIOR FILING DATE: 2000-04-06
! PRIOR APPLICATION NUMBER: DE 10019173.8
! PRIOR FILING DATE: 2000-04-07
! PRIOR APPLICATION NUMBER: DE 10032529.7
! PRIOR FILING DATE: 2000-06-30
! PRIOR APPLICATION NUMBER: DE 10043826.1
! PRIOR FILING DATE: 2000-09-01
! NUMBER OF SEQ ID NOS: 350
! Seq ID NO 206
! LENGTH: 11745
! TYPE: DNA
! ORGANISM: Artificial Sequence
! FEATURE:
! OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
! NAME/KEY: unsure
! LOCATION: (9105)
US-10-240-453-206
Query Match      3.1%, Score 145.4; DB 14; Length 11745;
Best Local Similarity 43.5%; Pred. No. 3.3e-13;
Matches 968; Conservative 0; Mismatches 1241; Indels 18; Gaps 6;

QY    328  ATTTCGAATATTACCATATCAGTACTAGAGTCTTTATCGACATCATTAATAATCAAT 387
DB    6029  ATTAAATATATCTAACAATATATACCTTAAATATATTAATTAATTAATTAATTAATTAAT 5970

QY    388  CAAAAAATAATATAGTACTATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 447
DB    5969  AATATATTAATTAATCCCTATATATCTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5910

QY    448  ACTCAATATTAATAGAGACCTTAACCTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 507
DB    5909  AAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5850

QY    508  TCATATTTAGCATATATGCATCTGCACATGATTAACATATGTAAACACATTTCTTAAAC 567
DB    5849  AATTAATTAATTAATCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5790

QY    568  TAGATTAATTAATCAATTAATATATATATTAATTTTGACCTGGTTCTGATATTAATTAAGTA 627
DB    5789  AAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5730

QY    628  AAAAAATGAATCAATTTTAAAGTGACGATCATTTGGCACAGCAATAGCAATTTGCACT 687
DB    5729  AAAAATAATTAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTTAAAT 5670

QY    688  GTCAAGCATAGTATATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 747
DB    5669  TTATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5610

QY    748  AAATCACTAACGAAAATTTAAATTAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 807
DB    5609  TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5550

QY    808  TGCAACCAAGCAATTAAGCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 867

```

Db 4484 TCTAACCTTTAAAAAATCTAAATPAAAAACAAAAACAAAAAATAAAAAATPAAATACCAT 4425

Qy 1346 AAAAGAGCAACAAATGCATATATTTAAATAGTAAAAACCTATTACTACGTAAATTTACAT 2005

Db 4424 TAAATATAAAAAATCTATACAAATTAATATCTAAATTAACATAAAAAATCCCTTTC 4365

Qy 2006 AGATGTCCTGACATATATAAATCTATATCTATGATCCTTGACATATTTCAACATATAATATGATA 2065

Db 4364 AAAACAATAAAAAATTAATCAATTAATCATATATACCTTTAAATCCTAATAAACAAAACTAAAA 4305

Qy 2066 AAGCTTTTTTTTANGCCATTTAGTTCTTAATPAAATGCTAAATTAACCCAAAAAGGTAAA 2125

Db 4304 ATCATTTAATTTTATCATATATATACCTCAATTAATTTAATTTAATTAATTAACAACTCTTTA 4245

Qy 2126 ATACACATTCCTGACACAAATATGAAAGATPACATATCAGCTCAATPAAATTAACAAAT 2185

Db 4244 TTAATAATTAATAAAAAATATATATCTATATTAATTAATTAATAATAATAATAATAAT 4185

Qy 2186 TTCTAATATACCCCTACTCTAAACATAGGAACCAATAGT-ATATCATGCTTAAGTAG 2244

Db 4184 AAAACACACAAATTAACACAACTTTAATAAAAAATTTCTCATATAAAAAACAATAAAA 4125

Qy 2245 ATATCCCTAGACAGATGATACAAATPATAAGAAATAATCGTTATGATATTTATACAAATTA 2304

Db 4124 TAAACGTTAATATAAAAAATATATAATCAAAAAAATTTCTTAAATTAATATATTTAAAA 4065

Qy 2305 CAAATTTCCAACTCAACAAACAAACTCAAGTCTATATATACAAAATCAAGAAATTA 2364

Db 4064 CAACTCATATATCAATTAATAATATACCAATTAATAAAAAAACHCAATATTTACAAACACA 4005

Qy 2365 ATATATCTTCCAAACACATATGATACAAATTTCTATAATATACATCATATCTTTTAAAAAG 2424

Db 4004 AAAAAAATCTTACAAAAAATTAATAATCTTTAAAAAACAATTAAAAATTAATAAAC 3945

Qy 2425 ATGACTTCACAAATCATACATAAATATTTATPACATGAAATGTTATTCAGAAACATTACA 2484

Db 3944 ACBAATTAATAAACTATACCTTAATTAATAAAAAAATCACTTTCCATTTTACAAAAAATA 3885

Qy 2485 CAATCCAGAAATTTTCAATTTTACGGTGTAATGCTGAACGTTCCATTCGAACAAAATTCAGC 2544

Db 3884 AAATATATTTATATTAATTTAAAAACAATTTACTAAAAATTAATTAATTAACAAAAAATA 3825

Qy 2545 AGTTCGT 2551

Db 3824 AATTCTT 3818

RESULT 28

US-10-311-455-986/c

Sequence 986, Application US/10311455

Publication No. US20030143606A1

GENERAL INFORMATION:

APPLICANT: OLEK, Alexander

APPLICANT: PLEBENROCK, Christian

APPLICANT: BEELIN, Kurt

TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determination of the Cytosine Methylation

TITLE OF INVENTION: Cytosine methylation

FILE REFERENCE: 5013.1014

CURRENT APPLICATION NUMBER: US/10/311.455

PRIOR FILING DATE: 2002-12-16

PRIOR APPLICATION NUMBER: PCT/EP01/07537

PRIOR FILING DATE: 2001-07-02

PRIOR APPLICATION NUMBER: DE 10032529.7

PRIOR FILING DATE: 2000-06-30

PRIOR APPLICATION NUMBER: DE 10043826.1

PRIOR FILING DATE: 2000-09-01

NUMBER OF SEQ ID NOS: 2424

SEQ ID NO 986

LENGTH: 7597

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

US-10-311-455-986

Query Match 3.0%; Score 144.2; DB 14; Length 7597;
Best Local Similarity 47.4%; Pred. No. 4.3e-13;
Matches 556; Conservative 0; Mismatches 613; Indels 4; Gaps 4;

1294 AATTAAACGCTATATGCAAAAATATTCATAGTCTTCTATATTAATTTGGCC 1353
1272 AAAACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1213
1354 ATCATGCTTGGAGATCTAATATTAACATGTAACAAGACATTCACCAATGGCTT 1413
1212 ATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1153
1414 TGGACTAGAAATGAGAAAGTAAATTAATTAATTAATTAATTAATTAATTA 1473
1152 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1093
1474 TGTGAAGAACAGTACTGTAATAACCATGATATCATAGCAAAAATTAATTA 1533
1092 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1013
1534 ACTACCAATTTGCTATCTATTTACATTTATTAATTAATTAATTAATTAAT 1593
1032 AATACATTAATCATTAATTAATTAATTAATTAATTAATTAATTAATTA 973
1594 CATATCAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1653
972 CATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 913
1654 AACGCTCAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1713
912 AATACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 853
1714 TGTGATCACTCACTAGTATTAATTAATTAATTAATTAATTAATTAATTA 1773
852 AATACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 793
1774 TTCTAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1832
792 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 733
1833 CGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1892
732 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 673
1893 TTTTCACTACATCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1952
672 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 613
1953 ACCAATCACTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2012
612 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 553
2013 CTGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2071
552 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 493
2072 TTTTATGCGCTTGTCTTAATTAATTAATTAATTAATTAATTAATTAATTA 2131
492 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 433
2132 ATTCCTGCAACATATGAAGATGATCAATCAATCAATCAATCAATCAATCA 2191
432 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 373
2192 TATTAACCACTACTTAATTAATTAATTAATTAATTAATTAATTAATTA 2250
372 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 313
2251 CTACACATGATCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2310
312 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 253

QY 2311 TTCCAACTTCACACAAAATCTCAAGTCTATTAATTAATTAATTAATTAATTA 2370
DB 252 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 194
QY 2371 TTCACCAACATAGTATTAATTAATTAATTAATTAATTAATTAATTAATTA 2430
DB 193 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 134
QY 2431 TCACAACTTCACATTAATTAATTAATTAATTAATTAATTAATTAATTA 2463
DB 133 AAAAAATACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 101

RESULT 29

US-10-312-841-2/c
; Sequence 2, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
; APPLICANT: EpiGenomics AG
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MH
; FILE REFERENCE: E01/1208/KO
; CURRENT APPLICATION NUMBER: US/10/312, 841
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 2
; LENGTH: 3673778
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: chemically created genomic DNA (Homo sapiens)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (379615)
; US-10-312-841-2

Query Match 3.0%; Score 141.6; DB 14; Length 3673778;
Best Local Similarity 43.8%; Pred. No. 1.2e-11;
Matches 1001; Conservative 0; Mismatches 1269; Indels 18; Gaps 8;

QY 374 ATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 433
DB 2092332 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2092273
QY 434 AGCTTTAGCAAGACCTCAATTAATTAATTAATTAATTAATTAATTAATTA 493
DB 2092272 CCACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2092213
QY 494 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 553
DB 2092212 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2092155
QY 554 ACATTTCTTAACCTAGTATTAATTAATTAATTAATTAATTAATTAATTA 613
DB 2092154 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2092095
QY 614 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 673
DB 2092094 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2092035
QY 674 ATGCAATTCAGCTGACGACATGATTAATTAATTAATTAATTAATTAATTA 733
DB 2092034 TTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2091975
QY 734 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 793
DB 2091974 ACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2091920
QY 794 CCAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 853
DB 2091919 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2091860
QY 854 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 913

Db	20918659	AAAAATCCAACCTTACATACATTAATAAAACAATTAATCAAAAAATTTAAAAACCCAAAC	20918800
Qy	914	CAGCATATGTGTAATAAATATCCCAATATTAATTTGTACTTAAGATCGAATATACATCA	973
Db	20917959	AATACTCTTCACTTATATATATTAATTAATTTATATAAACCAATACCTAATATATTCATTA	20917400
Qy	974	CT--CAATTTCCAGTGAATATAGCAGAGAAATTTTTCATATTAATCCAAATTTTATATA	1030
Db	20917339	AAAAAAAATATATCTTTTCATCAACAACATACAAAAACAACAACATACCTATACAAAA	20916880
Qy	1031	CTCTCGGCGCAAGTTTGTCTAAGAATGTGAGACATCTTCCCTGTGATGTAGTACTT	10950
Db	2091679	AAATTAATACTAATCCCTACTTCACACCAATATAAAAATTAACCTCAATATACAAAC	20916200
Qy	1091	GCTGGTGAATATATMAAGAACTTGGTGATCATTTGATAGAACATATAGTATGATTTCTA	1150
Db	2091619	CTAAATATATAAACTTAATATTTAAAAAACTCTTAATAAAAAAAACATTAACATATAACTTCAT	20915600
Qy	1151	AAAAATATATATCCATCAAGATATTTATAGGTGACAGATTGAGCGTGCATTAAGAACATA	1210
Db	2091559	AACCTTAATTTAATTAATAAAATCTCTAATCTTTAATATACAAAAACCAATACAAAAAA	20915000
Qy	1211	ATTGCAATGTCATGTGGAATATCGCTGGAATAATTTAGGTAAATAGTCTGTTGCTACT	1270
Db	2091499	AAAAATTAATAAATTAATTAATTAACATCATCAAAATTTAAAAATATATTAATCTAAAAACCA	20914400
Qy	1271	GTTATTAATCTAAGAGATATAGAAATTAACAACATATATATAGCAAAAAATCATTCAATA	1330
Db	2091439	TCAAAAAATAATAAAAAAATCTCATTTAAATAAAAAAATTTTACAAATCTTAATCTTA	20913800
Qy	1331	GATCTTCATACATTAATTTGTCATCATGTCCTTGAGATCTAATATTAATTAACATGTACA	1390
Db	2091379	CAAAAAAAATCTATATCTAATAAATATATATAAAAAATATTAATTAATCTCAATATTAATAT	20913200
Qy	1391	GAACTTACACGCAATATGTCCTTTTGACTAGAAATAGGAAGATGAAATATTAATACA	1450
Db	2091319	TATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	20912600
Qy	1451	TTAATATAGTCAACAACTTAAGCTTGTGTAAGAAACCACTACTGTAAACAATGATACA	1510
Db	2091259	AAACCCATTAACAACATAAAAACATATTCATCATCAATTAACATCAAAAAATTAACACATCA	20912000
Qy	1511	TTAGCAAAAAAATCTTAATGTGAACTACCAATTTGCATATCTATTTTCAATTTTTCAT	1570
Db	2091199	ACCACAAATTAATATATCTAT--TTCCACACCCCTAATATACCTATTAATTTAAAAATATA	20911410
Qy	1571	GAGAAATATATCATAGATTAAGCCATATCAACAATTTATCTGTAGCTATCATCTTTGA	1630
Db	2091140	ATTATATATTAATTAATAAATATTAATAATCTAATAATACACACTTTACTAATTAATAATTA	20910810
Qy	1631	TAAATTCACACAAATGATTAATAATTAACGCTCAAAATATAGATACAACGTTACACACATA	1690
Db	2091080	AAAAACATTAACCCCTTTAAAAAACAACCTAACAATACCTCAAAATTTTAAACATTAATTT	20910210
Qy	1691	TATAAAAAGCAGTCMAATATCCCATCGTTACATCCCACTAAGTATGTAACATATTACATA	1750
Db	2091020	ATTACATTAACCCAAATATCCCTCTCTTAATAA---CAATATACCCAAAAAATAAATAAC	20909640
Qy	1751	ACAAACTTAAGTATTAATCTTAAATTTTCCCTACTAATTAATTAACAATTCCTATCTGCT	1810
Db	2090963	ATTAATTCACATTAATAAATCTTATATACATATTTATTAACAACATTTATTAATCACAATCCAA	20909040
Qy	1811	ATCTCAATTAACAATTTGTATTCAGATTAATAATTAAGTCAAAATTAACCTTTTACTTAATTT	1870
Db	2090903	ATAAAAAATACCAAAATATTCACAACTATTAATAATTAATTAATTAATTCATATCAATTA	20908440
Qy	1871	ATAAATATATATGTGTGCACACTTTTTCAGTAC--ATCATGTATATATATGACCACTAATTT	1928
Db	2090843	ATATATTCACACGATTAATAAATAATTAATAATATTAATAATTAATAAATCTTACGACCCACAACT	20907840
Qy	1929	TACATTAATAATCTTAATTAATAAAGAACCAATATCATATATTAATTAAC--TAAAAACCTAAT	1987

```

Db      2090783  TTAAGAAATATTCGAGTAAATATAAAAAAACAACACAAAAACCAATATTATTAATTT 2090724
Qy      1988     TACAGACGAATATATACATAGTGTCTCGACATATAATTTCTATGACCTGTGACATATTT 2047
Db      2090723  CATTCATATAAAATATTCGAAATATAATATACTATATAAATTTAAACATATAATTTATTT 2090654
Qy      2048     CAAACATATAAATATGATTAACCTGTTTTTTTTATGCGCATTAAGTCTTATATAATTTGCTAA 2107
Db      2090663  CTAAAACTAAAAATTTAAATATTAATTTTCTCAAAACTAAAAAATAAAAAATCT 2090604
Qy      2108     TTACCCAAAAAGTTAAAAATACAGATCCGCAACAATATGAGAAGATACAAATACAGCT 2167
Db      2090603  ATAAATTAAATTTAAATATAAAAAACACACTATAATTTCTTAAAAACCACTAAATTTAT 2090544
Qy      2168     CAAATTAATTAACAAATTTCTAATATACCACTACTTAAACATAGAACCAATAGTA 2227
Db      2090543  ATTTTAAATATACAAATTTTATACATATTATTAATTAATATTCCAATTCCTTAAAAATAT 2090484
Qy      2228     TATCATACCTTAGTAAGATATATCCCTACAGCATGATCAAAATTAATAAGAAAAATCGTAT 2287
Db      2090483  TAAATTACATATAAACTTATATCATTTACACTCAACAAAACTACAAATATAAAACAC 2090424
Qy      2288     GTATATTATTAACAATTACAAATTTCCAACTTCACACAAAAAATCTCAAGTTCTATTATA 2347
Db      2090423  TCACATATTAATTTACGTACAAATATAAAAAAATTAACATACAAAAAATTAACATAATA 2090364
Qy      2348     ACAAAATTCAGAAATTAATATATATCTTCCACAAACATGTATATCAAAATTCATATAATAC 2407
Db      2090363  ATAAATCACCTTAATTAATTAATTAATTTAAATTTAAATAAAAAATTTAAATAATATATTT 2090304
Qy      2408     ATATCTTTTTTAAAAAGATGACTTGACAATTCATACATAAA-TAATTATACTAGCAATGT 2466
Db      2090303  ATACATATTTTAAATTAATTTATTTATTTCTCCATATTAACCAAAATTAATTTTACATAAAT 2090244
Qy      2467     TTATTCAGAAACATTAACAATCCAAAGAAATTTTCAATTTACGGGTAAATGCTGACGTTCC 2526
Db      2090283  TAAAAATTAATAATTTCAAAATCTTCAAACTCTCAATCAATTAATAACATATAACAAAT 2090184
Qy      2527     CATTCACAACAATTCAGACAGTTCGTGATACATTAAGCTCCAAAAGAGATACAGAGTTTT 2586
Db      2090183  AATATATAACTATCTAAAAATATCTTCTTAAAAAACTAACTAAAAACATATCATACAAAA 2090124
Qy      2587     TAAATACGTCTTCAATTCCTTACAGAGCTTACGTAAATGATACAGAAATTAAGCTTAGC 2646
Db      2090123  AAAAAACAATACTAAAAATTCAAAAACAAACTACAAATATATAAAAAAATAAAAATATAA 2090064
Qy      2647     ACTTAAAA 2654
Db      2090063  ACATAAAA 2090056

RESULT 30
US-10-398-221-3618
; Sequence 3618, Application US/10398221
; Publication No. US20040018514A1
; GENERAL INFORMATION:
; APPLICANT: KUNST, Frederick
; APPLICANT: GLASER, Philippe
; TITLE OF INVENTION: Listeria innocua, genome and applications
; FILE REFERENCE: 344 702 - US
; CURRENT APPLICATION NUMBER: US/10/398,221
; CURRENT FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: FR 00/12 697
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 4025
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3618
; LENGTH: 2791
; TYPE: DNA
; ORGANISM: Listeria monocytogenes 4b
; FEATURE:
/

```

```

; NAME/KEY: misc.feature
; LOCATION: (1)..(end)
; OTHER INFORMATION: n can be any nucleotide: a,g,c or t/u
US-10-398-221-3618

```

Query Match	2.9%	Score 139.8;	DB 15;	Length 2791;
Best Local Similarity	49.4%;	Pred. No. 1.5e-12;		
Matches 480; Conservative	0;	Mismatches 477;	Indels 14;	Gaps 4

QY	613	GATATAAATTAAGGTAAATAAAATGAAATACGATATTTTAGTGCAGGATCATTTGGCAGAC	672
Db	664	GAAAAATATTAAGACACAAAAAAAGTAGCTHTCTGGCCCTGGAAAGTTGGGAACAAG	723
QY	673	AATAGCAATTGACACTGTGCAGCATGTATATCAGTTAATCTTATGSSGACGTATATAG	732
Db	724	TCTTGACCTGCTCTAGCTGATTAATATATCAACCAAGTTATTTGGGGAACCTTAGATATA	783
QY	733	AAATATTAACCATATTAACACTTACCGAAAAAATTTAAATATTTACCCACATATACAT	792
Db	784	AATTGTGAATGAATATTAACGAATCGCACCGAATATGTCACTAATTTGCCAGATATTATTT	843
QY	793	ACCAGACAAATATATATGCAACCGACGATATAGACGAATATATCTGACACACATACATG	852
Db	844	ACCACTGAGGTAAAGCAGCATGTGTCACTTAGAAGCATATGATGTGTGACAGAAAT---	900
QY	853	TATATCTTACATATCTCTTACACAACAATTAGCACCATATGTACACAAATACAAACAA	912
Db	901	TGTGTATATTTGCTATTTCCAAACAAAGCAGTGGAAATGTTGTATAGCATTAATGAAGC	960
QY	913	ACAGCATATGTGTAAATAATCTCAATTAATTTGTATGAAGATATCGAAATTAATCATC	972
Db	961	GCTAAA-AGAACCAACATTTTATAGTACATGTAGATGAAGATATGAACGAAACAAATC	1011
QY	973	ACTCAAAATTTCCAG---TGAATAAGCAGAAAGAAATTTTACATATATCCAAATTTTAA	1022
Db	1020	TTGCAATGTCCGAAGATTTAGAACGAAATAGATGCGCTCCAAACGTAAAGCCCTTTGTG	1077
QY	1029	TACTCTGGTCCAAAGTTTGTCTAAAGAAATGCGAAACATCTTCTGTAGATATGATAC	1088
Db	1080	TTCTTTCTGACCAACATCTATGCGAAGAAAGTTGGCCTTCGTATCGACAAACACTTTGG	1133
QY	1089	TTGCTGTGTATTAATAAGAACTTGGTGAATCATGTATAGAAACAAATAGATATGATGTTG	1144
Db	1140	CGAGCTGTAAAGTTTATCAGCTGCTGAAATTTGTTCAAGATCGTTTCATCAATAAATAAT	1199
QY	1149	TAAAAATATATATCATCTCAAGATATTATAGGTGCACATTTGAGCGCTATTAAAGACA	1200
Db	1200	TGGCTATTTTATAGAAATGATATGTGATTTGGTGTGCAGAAATTTGGTGGCCCTGAAAAATA	1255
QY	1209	TAAATGCATTTGCATGTGGAATPATCGCTGAAAAAATTTAGTAAATATGCTGTGTCTA	1266
Db	1260	TTATTCGACATAGGTGCGAGAAATTTCTGACGGCTTGGTTATGGCGATATATGCTAAAGCGG	1311
QY	1269	CTGTATTAATTAAGCATGAAATGAAATTTAAACACTATATATGCAAAAAAATCATATCA	1322
Db	1320	CACATATGACTCGCCGGATGCGAATATCACTCGCTCGCTGTTC-----TGTGGCT	1377
QY	1329	TAGATCTTCATCATTATTTGGTTCATCATGCTTTGAGATCTATATTTAATACATGTACAA	1388
Db	1374	CTAATCCGCAAACTATTTACGATATGACAGGATTTGGATCTTATCGTTACTTGTACTA	1433
QY	1389	CAGAACTTACGCAATATGGCTTTTGTGACATAAGAAATTAGAAAAAGTATGAATAATAATA	1444
Db	1434	GTTGTCATTCAGTAAATTTGGCGTGTCTGTAAATATGCTATGAAAGGCAAGAAATTTAGATG	1499
QY	1449	CATTAAATAGTCACAACCTTAAAGCTTGTGTGAAGAACCAAGTACTGTAAACCACTGATAT	1500
Db	1494	AAGATATGAAAAAATGGGCATGTGTCTTTGAAGGTTCGAACACGTAAAGCAGTACATAG	1555
QY	1509	CATTAGAAAAAAATCTTAATGTAGAACTAACCAATTTGTATATCTATTTAAATTTATAC	1566
Db	1554	GTTGGGGAAAAACATAGATATTTGATATGACCAATTTACGAATGCATTTAAGCGCATTTAT	1611

QY	1569	ATGAGATATA	1579
Db	1614	TCGAAATATA	1624

```

RESULT 31
US-10-311-455-1280/c
; Sequence 1280, Application US/10311455
; Publication No. US20030143606a1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEBENROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Detecting Cytosine Methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1280
; LENGTH: 6175
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1280

```

Query Match	2.9%	Score 139;	DB 14;	Length 6175;
Best Local Similarity	45.8%	Pred. No. 2.7e-12;		
Matches 800; Conservative	0;	Mismatches 910;	Indels 35;	Gaps 8

Cy	740	ACACATCTAAACACTTACCGGAAAAATTAAAAATTTACCCACATATCATCTACGAC	799
Db	6002	ATTAATTTAAAAACAAAATATATATTAACACTTTAAATATACAAATTAATCTTAATAT	5943
Cy	800	AACATATATGCAACACGACCAATATAGACGAAGTATATCTGCAACAAATACATGATATTC	859
Db	5942	TATATATTAACAAACGATATAACATACAAATTAATATACATAAACAAATATATATACAT	5883
Cy	860	TTAACTATTCCTACACACAACTTACGACCAATATGTACACAAATACAAACATAACAGAT	919
Db	5882	ACGTAAACCAATATTAATAATCAAAATATTAATTAATTAATTAATTAATTAATTAATTAATTA	5822
Cy	920	ATGTGTAAAAATACTCCAAATTTATTTGTAGTAAGGATCGAAATTCATCACTCAAA	979
Db	5822	TAAATATPAAAAAACATATATATAACGATATAATTAATTAATTAACAAAA	5767
Cy	980	TTTCCCAGTGAAATAGCAGAAAGAAATTTTACATATATCCAAATTTTATCTCTCGT	1039
Db	5766	ATTAACATATATTAATAACAAATATATTAATATTAATAACATATATTAATAACATATATTA	5707
Cy	1040	CCAAGTTTGGTAAAGAAATTTGCAGAACATCTTCCTGTAGTATGACTTGCTGCTGAT	1099
Db	5706	TACAAATATATATAACAAATATATATAATTAATTAATTAATTAATAACATATATATAAA	5647
Cy	1100	AATTAAGAATCTGGTAATCATTGATAGAAACAATAAGTATGATGTTCTTAATAATATA	1159
Db	5646	TATTAATAATTAACATATATTAATAATTAACATATTAATTAATTAATTAATTAATTAATTA	5587
Cy	1160	TACCATCAAGATATTAATGAGTGTCAGATTTGGAAGCTCATTAAGAAACATTAATTCGAAT	1219
Db	5586	TAAATATATTAATAATTAACATATATTAATAACATAATTTATTAACATAAATATATATA	5527
Cy	1220	GCATGTGAATTAATCCGCTGAAAAAATTAGATATTAATGCTGTGTCATCTGATTAAT	1279
Db	5526	AAATCTTAATTAATAATATTAATAATAATATATATATTAATAATAACATATATATTAAC	5467

Db	14647	TAAATTAACCTTCTAAATTAATAAATAATTAATTAACATTTTCATTTAATTATTCGA	14568
Qy	900	AAATACACACAAACAGCATATGTGTAAATACTCCATATTTAATTGTAGTAAAGTGA	959
Db	14587	ATATATATTTTAATTAATTAATTAACACAAATATGTATTAATTTTCATPACAAATTCCTATTA	14528
Qy	960	TCGAATAATTCACCTCACTCAAAATTTCCAGTGAATAATAGCAGAGAAATTTTACATATATATC	1019
Db	14527	CTATTAACAACTGACAAACAAA-----ACATTTTCATTTACATTTTAACATTTTATA	14477
Qy	1020	CAATTTTATATCTCTGTGTCCAGTTTGTCTAAGAAATATGAGAAACCTTCCCTGTGA	1079
Db	14476	AAACCTTACCCCATTAATTAACCATTAATAAATTTCTACAAATTAATAAACTATATTTCTCT	14417
Qy	1080	GTATATGATCTGTGTGTATTAATAAGAACTGTGTATCATTTGATATGATGAACATTAATGTA	1139
Db	14416	ATATTATATTAATTTCTTAATAAATATATATTTCTAACAATACATTAATTAATAAATCTAATA	14355
Qy	1140	ATGATGTTCTTAATAATTAATATACATCAATATTAATAGTGTACGATTTGGAGCTGCAT	1199
Db	14356	ATTAAACCTATATTAATAACACACACATATATCTCAATAAATATATATATTTTCAAACTT	14257
Qy	1200	TAAAGAACATATATTCGCAATTCGATGTGGAATAATCGCTGGAATAAATTTAGATTAATATG	1259
Db	14296	TAAATATCATATACAAATTAATTAATTAATTAATAATTAATAATCATTAACAAAATTAATAATA	14237
Qy	1260	CTGTGTCTACTGTATTAACGA- AAGGATGAATGAATTTAAACACTATATATAGCAAA	1318
Db	14236	TTTAAATAATTAATTAATTAATTAATTAATTAATTTTAAACCTACAACTTAATAAATAATCTTCA	14177
Qy	1319	AATCATTCGAATAGATCTTCATATGATTAATTTGGTCCATCATGTCTTGAGATCTTAATATTA	1378
Db	14176	AATTTATTTTAATTAATAATTAATTAATTAATTAATTTTAAATAATTAATTAATTAATTTATTC	14117
Qy	1379	ACATGTACACAGAACCTTACGCGCATATGGCTTTTGGACTAGAAATAGGAATAAGGTAGA	1438
Db	14116	TAAACAAACAACTTAATAAATTAACAAAT- - -TTATTAATAATAAATAAATAAATAAATA	14062
Qy	1439	AATTAATAATCATTAATTAATGATCAACAACCTTAAGCTTGTGAAGAACAGTACTGTAAAA	1498
Db	14061	AAACCAATTAATAAATAATTAATTCATTCMAAAATTAATTTTAAACCTAATAACATACGA	14002
Qy	1499	CCACTGATTCATTAAGCAAAAAAAGCTTAATGTGAACATACCAATTTGCATATCTATTTC	1558
Db	14001	ATCATATTAACCTATTAATAAATAAATAAATAATTAATTCACAAATCTAATTAATAAATAAATA	13942
Qy	1559	AATTTATTAATGATGAATATATATCTGATTAAGCAATATCAAA- - -CATATATCTTAG	1615
Db	13941	AATTAATAATTAATAAATTTTAATTCATTAACAAATCTAAATTAATTTCAATTAATTAATTAAT	13882
Qy	1516	TCTATCATCTTGTATTAATTTCAACACATATGATTAATAAAGGTCAAAAATATA- -GAAT	1673
Db	13881	TTAAAAACCAAAATTAATAATTTTCTAAAAAATAAATAATTAATCAAAATTTATTTATA	13822
Qy	1674	ACAACTGTCAACCATATATTAATAAAGCAGTCAAAATCCATCGTTATACATCCAACTTAAT	1733
Db	13821	AAAAAATTAATAAAGCTTTAAAAAATATATATACATTAATAATTAATTAATTAATTAATTAATTA	13762
Qy	1734	ATGTATACATTAACATTAACAACTTAAGTAACTTAATAATTTCCAGCTAATTAATATA	1793
Db	13761	TCCCACTCTCAAAAAACACAAACCTAAATATATTTCAACAACAAAATTTTACCAAACTTGA	13702
Qy	1794	ACAAATTCCTATCTCTATCTCAATTAACCAATTTGTATTCAGTTTAATAATTAAGTCAAT	1853
Db	13701	AACATATTAATATCTATCTTATATTAATTAATTTTAATTAATAAATAAATAAATAAATAAATC	13642
Qy	1854	TAACTTTTACTATTAATTAATAATAATATGTGGCACTTTTCTCACTCATATGATATA	1913
Db	13641	TACCAATATATTTTAATAACAAATTAATTTTAACATCAAAAACAAAAAATAAATAA	13582
Qy	1914	TAAATGAACATTAATTTACATTAATAATACTAATAAAGAACCAATACATATATTA- - -	1969

Db	13581	TTAACTATATATAGTCTCCAAAAACAATTATTAATATTAATTAATTAATTAATTAACAT	13522
Qy	1970	-----ATACTAATAACAGCTATTACTAGGTATATATACATAGAGTCCCT	2014
Db	13521	AATTCACAATTTAT	13462
Qy	2015	GACATATATAAATTCATTCATCCTTGCATATTTCCAAACATATAAATATATATATAT	2074
Db	13461	AAAAAT	13402
Qy	2075	TTTATGCCCATTAGTTCCTTAATTAATTCGTATTTCCAAAAAGTATTAATACACT	2134
Db	13401	CCATTCCTATTTTAAAAACCTATATATATATATATATATATATATATATATATAT	13342
Qy	2135	CTGCACAACATATATGAAGATATCAATACCGTCAATTAATTAACAAATTTCTATAT	2194
Db	13341	AATTAATTTTAAAAATATATATATCAACAATATATATATATATATATATATATAT	13282
Qy	2195	AACCACTATCTTAAACATATGAGAACCAATATATATATATATATATATATATAT	2254
Db	13381	AAATTAATATTCCTAT	13222
Qy	2255	AGCATATATCAATTAATTAAGAAAAATCGTATATATATATATATATATATATAT	2314
Db	13221	CACACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	13162
Qy	2315	AACCTCAACAACAACCTCAAGTCTATATATATATATATATATATATATATATAT	2374
Db	13161	ACAAAATTAACAT	13102
Qy	2375	CCAA--CACATAGTATCAAAATTCATATATATATATATATATATATATATATAT	2432
Db	13101	ACATTTTAAAAATTAATTAACAATTAATTAATTAATTAATTAATTAATTAATTA	13042
Qy	2433	ACAATTCAT	2492
Db	13041	ATATCTTAT	12982
Qy	2493	GAATTTTCAATTTAGCGGTATATATATATATATATATATATATATATATATAT	2552
Db	12981	CTATCCCCGTTTCCAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	12924
Qy	2553	ATACATTAATCCCTCCAAAACGATATATATATATATATATATATATATATATAT	2612
Db	12923	AAACACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	12864
Qy	2613	GAGCTTATAGTAT	2672
Db	12863	TATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	12804
Qy	2673	TAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	2732
Db	12803	TAAACAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	12744
Qy	2733	TTAAATTCATTTCTATATATATATATATATATATATATATATATATATATAT	2765
Db	12743	ATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	12711

RESULT 33
US-09-815-242-6968

Patent No. US20020061569A1

! GENERAL INFORMATION:

APPLICANT: Ohlson, Kari L.

APPLICANT: Zyskind, Judith W.

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John D.

APPLICANT: Carr, Grant J.

APPLICANT: Yamamoto, Robert

APPLICANT: AL, H. HOWARD
TITLE OF INVENTION: IDENTIFICATION

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

```

/ TITLE OF INVENTION: Prokaryotes
/ FILE REFERENCE: ELITRA.011A
/ CURRENT APPLICATION NUMBER: US/09/815,242
/ CURRENT FILING DATE: 2001-03-21
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ NUMBER OF SEQ ID NOS: 14110
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO: 6968
/ LENGTH: 1008
/ TYPE: DNA
/ ORGANISM: Haemophilus influenzae
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1)...(1008)
/ US-09-815-242-6968

```

```

Query Match      2.9%; Score 136.8; DB 9; Length 1008;
Best Local Similarity 50.4%; Pred. No. 2.9e-12;
Matches 476; Conservative 0; Mismatches 447; Indels 21; Gaps 5;

```

```

QY 620 ATTAGGTAAAAAAGAAATCAGTATTTTACGTCGAGATATTTGGCAGCAGCAATGCA 679
DB 4 ATAAGCTGGCAAAACCCATCAGTGTAGTGGGATCTTATGAACTGGCGCTGGCA 63
QY 680 ATTGACGTGAGCAGATGATATTCAGTTACTTATGGGACGTGATCATAAGAAATAT 719
DB 64 ATTACCTTTTCGGAATGTTCTCCCACTCATTTAGGGGCAATCCCGCCCATATC 123
QY 740 AOCATATAAACCTTACCGAAAAATTTAAATATTTTACCACATATCATCTTACGAC 799
DB 124 GCACAAATGCGAGAGAGACGAAATATGTTTGGCGGATGTCATCTTCCGAG 183
QY 800 AACATATATGACAGCAATATATAGCAAGTATATCTGACCAACATATCATGATATATC 859
DB 184 GATCTTCCTTTAGAAAGTATCTTGACACAGCAAT---GGATATTTCTCAAGATATTTA 240
QY 860 TTAACATATCTTACACACAAATTTAGCAGCATATGTAACAAATATACAAACAGCAT 919
DB 241 ATCGGGGCGCAACGCAATGCTTTCGGTGAATTTCTTATAAAAAT-----CAACCGCAC 294
QY 920 ATGTGTAATAATCTCCCAATATTAATTTAGTAAAGTATCGAA---ATTACATCATC 976
DB 295 TTTAAAGGCCCATCACCCATTTATTTGGGCAACAAAGGTTTGAACGTAAACAGGGCGT 354
QY 977 AAATTTCCAGTGAATATGCAAGAAATTTTACATATATATCCAAATTTTATATCTCT 1036
DB 355 TTACATCAAAACCGTAGTGAAGAACACTCGGAACGGAATATCCAGCTGATCTTCT 414
QY 1037 GGTCCAGTTTGTGTAAGAAATTTGCAAGACATCTTCTTGATATAGTATCTTGCTGCT 1096
DB 415 GGGCCAACTTTTGCAAAAGAAATTTAGCGAAGTGTACATCATAGTATTAAGCTTGCGCC 474
QY 1097 GATATATAAGAACTTG---GTGAATCATTTGATAGAAACAAATTAAGTATGTTCTAAA 1153
DB 475 AATAACGAGCAATTCGACGGGAATTCAGTCTGATTCATGTTATGATTAAGGTTTCCGA 534
QY 1154 ATAATATACCATCAAGTATTTATAGTGTACAGATTTGAGCTGCTATTAAGAACATTAAT 1213
DB 535 GTTATATATAATTTGGATATGACTGGCGTTCAACTTGTGTGAGCAATTAATAATGTGATC 594

```

```

QY 1214 GCATTCATGTGGAATAATTCGTGGAATAATTTAGTAAATAGCTGTGTACTGTT 1273
DB 595 GCATTCGTGGCGGTATTTTCAACAGGTATAGGATTTGGGCAAAATGCTGCACAGACTG 654
QY 1274 ATACTAAAGCAATGAATGAAATTTAAACACTATATATAGCAAAAAATCATTCATAGAT 1333
DB 655 ATACTCGAGGTATGCAAAATTAACCCCTTAGGTATCTC-----GCTCGGTGCAAT 708
QY 1334 CTTCATACATTAATGTCATCATGTCCTTGAGATCTCTAATTTAATAGTACACAGCA 1393
DB 709 ACCAATACATTTATGGAATGCTGTGATTTGGCGCATTTAGTCTTAATCTTGACCCGAT 768
QY 1394 CATTCAGCAATATGCTTTTGGACTAGAAATAGAAAAAGTAGAAATATTAATACATTA 1453
DB 769 CATCGCGCATGCAAGTTCGATTAATGTAAGTAAAGATTAAGATCCCAATGCT 828
QY 1454 ATAGATCAACACTAAGCTTGTGAGAGACCACTGTAACCACTGATATCATTA 1513
DB 829 ATGAAAAATATGCGCAAGTGTGAAATGCGATTTATATACAAAAAGCCATATTTGCTT 888
QY 1514 GCAAAAAAATTATATGATAGACTACCAATTTGCAATATCATTTA 1557
DB 889 GCACAAAGCAAGGTGTGAAATGCGATTTACAGAACAAATTTA 932

RESULT 34
US-10-282-122A-22052
/ Sequence 22052, Application US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari
/ APPLICANT: Zyskind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA.034A
/ CURRENT APPLICATION NUMBER: US/10/282,122A
/ PRIOR FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO: 22052
/ LENGTH: 1008
/ TYPE: DNA
/ ORGANISM: Haemophilus influenzae

```



```
FEATURE:
NAME/KEY: misc_feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (10091)..(10091)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (11924)..(11924)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
```

```
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (133160)..(133160)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152530)..(152530)
```

Query Match 2.9% Score 136.8; DB 14; Length 1830121;
Best Local Similarity 50.4%; Pred. No. 5.4e-11;
Matches 476; Conservative 0; Mismatches 447; Indels 21; Gaps 5;

```
QY 620 ATTAGTAAATAATGAAATGATATTTAGTGCAGAGCATTTGGCAGCAATAGCA 679
DB 635171 ATAACTTGCAAAACCAATCACTGTGCTAGTGGGATCTTATGAACTGGCTGCA 635230
QY 680 ATTGACGTGACGACGACGATATATCACTTAATTAGGGGACGATGATCAATGAATTT 739
DB 635231 ATTACTTTTCTCGCAATGTTCTTCCACTCATATGAGGGGCAAAATCCGCCCATATTC 635290
QY 740 ACACATATTAACACTTACCGAAAAATTTAAATATTTACCCACATATCATCTACAGAC 799
DB 635291 GCACAAATGACAGACGAAACGACAAATATATGCTTTTGGCGGATGTCATCTTCGAG 635350
QY 800 AACATATATGCAACGACGAAATATGACGAGATATATCTGACAAACATATCATCTATATC 859
```

Db 635351 GATCTTCATTGGAAGTATGTTGCAACAGCAAT---GGAATATCTTCAAGATATTTTA 635407
Qy 860 TTAATCTATTCCTACACAAATTTAGCAACCATATGTATACAAATACAAACAGCAT 919
Db 635408 ATCGTGTGCGCAAGCATATGCTTTCGTGAAATCTTATTAATAAT-----CAACCGGAC 635461
Qy 920 ATGTGTAAAAATATCTCCAAATATTTATTTGTATAGTATCGAA---ATTACATCATC 976
Db 635462 TTTAAAGCCCATACCGATTAATTTGGGCAACAAAGTTGGAGCTAACACAGGGCGT 635521
Qy 977 AAATTTCCAGTGAATATAGCAGAGAAATTTTACATATTAATTCATTTTATCTCTCT 1036
Db 635522 TTACATCAAAACCGTATGTTGAAGAACACTCGGAACGCAATTCACATGCTATCTTCT 635581
Qy 1037 GGTCCAGTTTGTCTAAAGAAATTCAGAACATCTTCTTGTATATGTATCTTGTCT 1096
Db 635582 GGCCCAACTTTTGCAAAAGAAATTAAGCGAGGCTTACCATGACTATTAACGCTTGCC 635641
Qy 1097 GATTAATTAAGAACTTG---GTGAATCATTTGATAGAAACATATGATGTTCTAAA 1153
Db 635642 AATTAAGAGCAATTCGACCGGAATTTCACTCTGATTAATGATAGTAAAGTTTCCGA 635701
Qy 1154 ATTAATATACATCAAGATTTATAGTGTACAGATTCGACTGATTAAGACATATTT 1213
Db 635702 GTTTATATTAATTCGAGATATGACTGCGCTCAACTGTGTGAGCAATTAATAATGTATC 635761
Qy 1214 GCAATTCGATGTGAATATTCGCTGAGAAATTTAGTAAATTAATGCTGTTACTGTT 1273
Db 635762 GCAATTTGGCGGGATTTTCAACGCTATGAGATTTGGCGCAATTCCTCGACAGCATG 635821
Qy 1274 ATTAATTAAGCATGAATGAATTAACACTATATATAGCAAAATTCATCAATGAT 1333
Db 635822 ATTACTCGAGTATTCAGAAATTAACCGCTTATGATCTC-----GCTGCGTCAAT 635875
Qy 1334 CTTTACATTAATTTGTCATCATGTCGTGAGATCTAATATTAACATGTACACAGAA 1393
Db 635876 ACCAATCATTTATGTGAATGCTGTGATTTGGCGATTTAGTGTCTGACCGATAT 635935
Qy 1394 CATTCACGCAATATGCTTTGACTAGAAATAGGAAAGGTAGAAATATTAATCATTA 1453
Db 635936 CAATGCGCAATTCAGCTTTTGAATTAATGCTAGTAAGATTCGATGCCAATGCT 635995
Qy 1454 ATAGATCAACACTTAAGCTTTGTAGAGAACCACTACTGTAAACGATGATCATTA 1513
Db 635996 ATGGAAATATTCGGCAAGCTCTAGAAAGTTTATTAATCAAAAGAACCTATTTGCTT 636055
Qy 1514 GCAAAAAAATTATGTAGAACTTACCAATTTGCATATCTATTTA 1557
Db 636056 GCACAAAGACAGTGTGGAATGCCGATTCAGAAACAAATTTA 636099

RESULT 36
US-10-329-670-1
Sequence 1, Application US/10329670
Publication No. US20040018503A1
GENERAL INFORMATION:
APPLICANT: Fleischmann et al.
TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Frag
FILE REFERENCE: P186P1
CURRENT APPLICATION NUMBER: US/10/329,670
PRIOR FILING DATE: 2002-12-24
PRIOR APPLICATION NUMBER: US 09/643,990
PRIOR FILING DATE: 2000-08-23
PRIOR APPLICATION NUMBER: US 08/487,429
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: US 08/426,787
PRIOR FILING DATE: 1995-04-21
NUMBER OF SEQ ID NOS: 1
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1830121
TYPE: DNA

ORGANISM: Haemophilus influenzae
FEATURE:
NAME/KEY: misc_feature
LOCATION: (4747)..(4747)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (9921)..(9921)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (10150)..(10150)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (29298)..(29298)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36543)..(36543)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36551)..(36551)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (440810)..(440810)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44975)..(44975)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44593)..(44593)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a, t, g or c

```
FEATURE:
NAME/KEY: misc feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (133340)..(133340)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (13360)..(13360)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
```

```
NAME/KEY: misc feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (152530)..(152530)
```

Query Match 2.9%; Score 136.8; DB 15; Length 1830121;
Best Local Similarity 50.4%; Pred. No. 5.4e-11;
Matches 476; Conservative 0; Mismatches 447; Indels 21; Gaps 5;

```
QY 620 ATTAGTAAATAAATGAAATGATATTTAGTGCAGATCTTTGGCAGCAATAGCA 679
DB 635171 ATTAAGTAAATAAATGAAATGATATTTAGTGCAGATCTTTGGCAGCAATAGCA 679
QY 680 ATTAGTAAATAAATGAAATGATATTTAGTGCAGATCTTTGGCAGCAATAGCA 739
DB 635231 ATTAAGTAAATAAATGAAATGATATTTAGTGCAGATCTTTGGCAGCAATAGCA 739
QY 740 ACACATATTAACATTTACGAAATTTAAATTTTACCAATATCATCTTACCAAT 799
DB 635291 ACACATATTAACATTTACGAAATTTAAATTTTACCAATATCATCTTACCAAT 799
QY 800 AACATATATGACACGACATATATGACGATATATGACGATATATGACGATATATG 859
DB 635351 AACATATATGACACGACATATATGACGATATATGACGATATATGACGATATATG 859
QY 860 TTAATATATGACACGACATATATGACGATATATGACGATATATGACGATATATG 919
DB 635408 TTAATATATGACACGACATATATGACGATATATGACGATATATGACGATATATG 919
QY 920 ATGTGTAATAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAA 976
DB 635462 ATGTGTAATAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAA 976
QY 977 AAATTTCCAGTAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAA 1036
DB 635522 AAATTTCCAGTAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAA 1036
QY 1037 GGCCCAAGTTTGGCAAGTAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAA 1096
DB 635582 GGCCCAAGTTTGGCAAGTAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAA 1096
QY 1097 GATATTAAGAACTGTG---GTGAAATCATGATGAAATGAAATGAAATGAAATGAAATGAA 1153
DB 635641 GATATTAAGAACTGTG---GTGAAATCATGATGAAATGAAATGAAATGAAATGAAATGAAATGAA 1153
```

Db 635642 AATAAGAGCAATTCGACGGGAATTGAGTCTGATTCATTAGTAAGAGTTTCCGA 635701
QY 1154 ATAAATACCAATCAAGATTTATAGTGTACAGATTGAGCTGCATTAAGACATTAAT 1213
Db 635702 GTTATATAAATTCGATATGAGCTGCGCTTCACTGTGAGCAATTAAGAGTATC 635761
QY 1214 GCAATTCGATGATGATTAATCCCTGGAAAAAATTTAGGTAATATAGTGTGTACTGT 1273
Db 635762 GCAATTCGATGATGATTAATCCCTGGAAAAAATTTAGGTAATATAGTGTGTACTGT 635821
QY 1274 ATAAATACCAATCAAGATTTATAGTGTGTACTGTGTACTGTGTACTGT 1333
Db 635822 ATAAATACCAATCAAGATTTATAGTGTGTACTGTGTACTGTGTACTGT 635875
QY 1334 CTTCATACATTAATAGTGTGTACTGTGTACTGTGTACTGTGTACTGTGTACTGT 1393
Db 635876 ACCAATACATTAATAGTGTGTACTGTGTACTGTGTACTGTGTACTGTGTACTGT 635935
QY 1394 CATTCACGCAATGAGCTTTTGTACTGTAGAAATAGAAAAAGTAGAAATATATATATTA 1453
Db 635936 CAATCCGCAATGAGCTTTTGTACTGTAGAAATAGAAAAAGTAGAAATATATATATTA 635995
QY 1454 ATAAATACCAATCAAGATTTATAGTGTGTACTGTGTACTGTGTACTGTGTACTGT 1513
Db 635996 ATAAATACCAATCAAGATTTATAGTGTGTACTGTGTACTGTGTACTGTGTACTGT 636055
QY 1514 GCAAAAAAATTAATGATAGAACTTACCAATTTGCAATATATATATTA 1557
Db 636056 GCAAAAAAATTAATGATAGAACTTACCAATTTGCAATATATATATTA 636099

RESULT 37

US-10-311-455-2213/c
Sequence 2213, Application US/10311455
Publicatation No. US20030143606A1
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPENBROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining the Cytosine Methylation
FILE REFERENCE: 5013.1014
CURRENT APPLICATION NUMBER: US/10/311,455
CURRENT FILING DATE: 2002-12-16
PRIOR APPLICATION NUMBER: PCT/EP01/07537
PRIOR FILING DATE: 2001-07-02
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 2424
SEQ ID NO 2213
LENGTH: 11691
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-2213

Query Match 2.9%; Score 136.2; DB 14; Length 11691;
Best Local Similarity 43.1%; Pred No. 9.5e-12;
Matches 1193; Conservative 0; Mismatches 1543; Indels 31; Gaps 10;

QY 256 ATAAATACCAATCAAGATTTATAGTGTGTACTGTGTACTGTGTACTGTGTACTGT 315
Db 11118 ATAAATACCAATCAAGATTTATAGTGTGTACTGTGTACTGTGTACTGTGTACTGT 11059
QY 316 TGTTCATCATATATTTTCAATTAATATATATATATATATATATATATATATATATATAT 375
Db 11058 TTTTCTTAAAT 10999
QY 376 ATAAATACCAATCAAGATTTATAGTGTGTACTGTGTACTGTGTACTGTGTACTGT 435

Db 10998 ATAAATACCAATCAAGATTTATAGTGTGTACTGTGTACTGTGTACTGTGTACTGT 10939
QY 436 TCTTGAAGAACTCAAT 495
Db 10938 AAAT 10879
QY 496 TTAATACCAATCAAGATTTATAGTGTGTACTGTGTACTGTGTACTGTGTACTGTGTACTGT 555
Db 10878 AAAT 10819
QY 556 ATTCCTTAACTAGATTAAT 615
Db 10818 AAAT 10759
QY 616 ATAAATACCAATCAAGATTTATAGTGTGTACTGTGTACTGTGTACTGTGTACTGTGTACTGT 675
Db 10758 CTCTTATA-AAAAAT 10700
QY 676 AGCAATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 735
Db 10699 AAAT 10640
QY 736 TATTAACAT 795
Db 10639 AAT 10580
QY 796 AGCAATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 855
Db 10579 AAT 10521
QY 856 TATCTTAT 915
Db 10520 AAAT 10461
QY 916 GCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 975
Db 10460 AT 10401
QY 976 CAAATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1035
Db 10400 TAT 10341
QY 1036 TGAT 10295
Db 10340 AAAT 10283
QY 1096 TGAT 1155
Db 10282 AAAT 10223
QY 1156 AATATACCAATCAAGATTTATAGTGTGTACTGTGTACTGTGTACTGTGTACTGTGTACTGT 1215
Db 10222 TTAACATTTTCACTAT 10163
QY 1216 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1275
Db 10162 ATCTCACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 10103
QY 1276 AATTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1335
Db 10102 ATCTCACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 10043
QY 1336 TCAATACATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1390
Db 10042 TAAATACCAATCAAGATTTATAGTGTGTACTGTGTACTGTGTACTGTGTACTGTGTACTGT 9983
QY 1391 GAATATTCAGCAAT 1450
Db 9982 TAAAT 9923
QY 1451 TTAATAGATCAATCAAGATTTATAGTGTGTACTGTGTACTGTGTACTGTGTACTGTGTACTGT 1510
Db 9922 ATAAATACCAATCAAGATTTATAGTGTGTACTGTGTACTGTGTACTGTGTACTGTGTACTGT 9863

[illegible][illegible]

APPLICANT: OLEX, Alexander
APPLICANT: PIERENROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Detect
FILE REFERENCE: 5013.1014
CURRENT APPLICATION NUMBER: US/10/311.455
CURRENT FILING DATE: 2002-12-16
PRIOR APPLICATION NUMBER: PCT/EP01/07537
PRIOR FILING DATE: 2001-07-02
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 2424
SEQ ID NO 1460
LENGTH: 6831
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1460
Query Match
Best Local Similarity 2.9%; Score 135.8; DB 14; Length 6831;
Matches 1808; Conservative 0; Mismatches 2322; Indels 89; Gaps 19;
239 AAGAACTCCCTCTCTCTGTTCCATCCATATTTTCATTTTACATATTCATAGATT 358
Db AATTAACCTTATATATATATTTCTATCTTAAATCAATCAAAAAAATAATTTT 6521
Qy CTTATTCGACATCATATTTAAATCATCAAAATAATATGT-----ACATTC 410
Db 6520 ATTCATTAACAAATATATCTTAATATCTTAATATCTAATATTTTACATTTCTAATA 6461
Qy 411 TAATTAATTAATTAATCTCTCATATGCTTTAGCAAGAACTCAATATTAATGAGACATT 470
Db 6460 CCAAAAATTAACATATTAACAAACCTTACCTCATCAATATTAATTAATTAATTAATTA 6401
Qy 471 AACTTATTAATTAATTTCAATTTTATTAATCTACTCATATTTAGCATATATTCAT 530
Db 6400 AATTAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6341
Qy 531 CTGACATGATTAACATATGTAACATTTCTTAACATGATATATATATATATA 590
Db 6340 AATATTAATTAATTAATTTCTTAATTAATTAATTAATTAATTAATTAATTAATTA 6281
Qy 591 TATATTTTGAATGTTCTTGAATATTAATTAATTAATTAATTAATTAATTAATTAAT 650
Db 6280 AATTAATTTTAA-----TACTAATTTTAAATTAATTAATTAATTAATTAATTAAT 6225
Qy 651 GTGCGATGATTTGGCAGCATATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 710
Db 6224 AAAAAAATTAATGATATGTAATCTTAACGAAATCTCTTAATTAATTAATTAATTAAT 6165
Qy 711 ACTATGAGGAGTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 770
Db 6164 TCTCTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6108
Qy 771 AATATTTTACCATATTCATATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 830
Db 6107 AAAAAATATATATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 6048
Qy 831 TATATTCGACATATGATATTAATTTTAACTATTTCTTACATCAATTAATGAGCA 890
Db 6047 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5988
Qy 891 TATGTAACCAATATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 950
Db 5987 ATTTTAATCAATTAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5928
Qy 951 GTAAGGATGAAATTAATCAATC-----TCNAATTTCCAGTGAATTAATGAGAG 1002

Db 5927 ACTAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5868
Qy 1003 AATTTTACATATATATCAATTTTATTAATTAATTAATTAATTAATTAATTAATTAAT 1062
Db 5867 ATTTAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5808
Qy 1063 AGAATCTCTCTGATATATGATGATGATGATGATGATGATGATGATGATGATGAT 1122
Db 5807 AGCTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5748
Qy 1123 GATGAACAT 1182
Db 5747 AACTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5688
Qy 1183 ACATGATGAGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1242
Db 5687 ATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5628
Qy 1243 AATTTAGTAAATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1301
Db 5627 ATTAATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5568
Qy 1302 CACTATATATAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1360
Db 5567 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5508
Qy 1361 CTGGAGATCTAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1420
Db 5507 ATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5448
Qy 1421 GAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1477
Db 5447 TATCTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5388
Qy 1478 GAAGGACATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1537
Db 5387 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5328
Qy 1538 CCAATTTGATATCTATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1597
Db 5327 CTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5268
Qy 1598 TCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1637
Db 5267 TCTAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5208
Qy 1638 AACATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1697
Db 5207 ACTACATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5148
Qy 1698 AGCAGTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1757
Db 5147 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5088
Qy 1758 TAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1817
Db 5087 TCAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5028
Qy 1818 TTAACCAATTTTATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1877
Db 5027 TTCCCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4969
Qy 1878 AATATGTTGACATTTTATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1937
Db 4968 AATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4909
Qy 1938 TAACTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1997
Db 4908 CTCAAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4849
Qy 1998 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2057
Db 4848 AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4789

QY	2058	CTATGATGAAAGCTGTTTTTTTATGCCCATTAGTCTTAATTAATAATGCTAATATACGAAA	2117
Dp	4768	ACATTACGACTATCTCTAATCTAATAATATATTAATAAATAAATAAATAAATCAT	4729
QY	2118	AGSTTAATAATACACATTTCTCTGCAACAAATGAGAGATACAAATACGCTCAATTAAT	2177
Dp	4728	ATTTTCTAATMAA---AGTATMAAATAAACAATAATACCTATATATTAATACAT	4673
QY	2178	AAACAAATTTCTAATATAACCACTACTCTAAACATAGAACCAATAGTATATCATCT	2237
Dp	4672	AAAAAAAAAATAAATAAATAAACAAGAAATMA---AGAAATTAATTTCTAATACT	4616
QY	2238	AACTAAGATATCCCTACAGCATGTATCAATATATAAGAAAAATCGTTATGTATATTAT	2297
Dp	4615	AAAAACCACTAAGCTATTTTATTTAACAATAAATAATATACATCAAAATCACCTCAAC	4556
QY	2238	ACAAATTAACAAATTTCCAACTTCAACAACAACAAACCTCAAGTCTATTTAACAATTCAA	2357
Dp	4555	AAACTTAATACMAAATAAATAAATAAATAATTAATAAACCATCGAATAATCTCTACCACT	4496
QY	2358	GAAATTAATATATACCTACCAACACATAGTATCAATTCCTAATACATCATCTCTTT	2417
Dp	4495	TACTTTTATAAACAACCTTTATTTTCATTCACACTTTCTCACCACTTCTAATAC	4436
QY	2418	TAAAGAATGACTTCACAAATCATATCAATAAATAATTAACANGAAGTTATTCAGAAA	2477
Dp	4435	AAAAATAAATAATTTAATAAT---TATTTAATACATTTAACAATAATAATAAATAA	4379
QY	2478	CATTCACAAATCCAGAAATTTCAATTAACGGTGAATGCTGAACGTTCCATCCAA	2537
Dp	4378	ATTCATTCCTTAACCTTTACCAAAATAATATCATATCTAATTTTAAATCTCATTT	4319
QY	2538	ATTGACAGTTCGATACATTTACTTACCTCCAAACAGATACTGATTTTAAATACGTC	2587
Dp	4318	AAATAATTAATAAATACTACTTAATAAATAAATAATTTATTAATAAACAATAAATAAT	4259
QY	2588	TTTCAATTTCTACAGACTTTAGTAATGGTACAGAAATACCTTTAGCACTTAACAT	2657
Dp	4258	ATATAAATAACAAATATTTTATTTCTCAACTCAAAAAAATACTTTAATAAATAATCTAT	4199
QY	2658	CATTAATAGGGAATTAATGATCTAATCTACTCTGTAATCTTTGCTGAATCATCTCCT	2717
Dp	4198	ATTATCTGTATATACATAAATAATCTAATAAACAATATATCATATCAATAAAGCTTAATA	4139
QY	2718	TAGGGAATTCATTAATTAATCCATTTATTAACATTAACGTAATGCTTAAGTTA--CC	2775
Dp	4138	TAAATAAATAATTTATTTTCTACCCATTAATAAACAATCTTACTATATATTTAATAA	4079
QY	2776	TGGCCATTCATATGACTCATTTGCTATTAAGCTTCACTTCACTTAATACATGATACAA	2835
Dp	4078	TAAATAAATAATTAATCTCATCATATAATAATAATAATAAATAAATAAATAAATAA	4019
QY	2836	ACCAATTTTTTACGAGATCTATTAATAATACCTACACAATCCGGTATATCTGTACA	2895
Dp	4018	ACATTTATCAAAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	3959
QY	2896	ATATTCTACTAAGACGGTACTCTAATTTGAGAGACATTTATCTATATAATAATCTC	2955
Dp	3958	AAAAATTAATACAACTCTAATAATAACATAATCAAAATTAATCTTAATATCAAAATAC	3899
QY	2956	ACAAACCTACCAAGTTTACTCTCACTTCAATCTTTGGAAGAACACATAATCT	3015
Dp	3898	TATAATATTAACATACCTTAACAATCATTTAATAATTTTACTATCTCTAATTAATAA	3839
QY	3016	CACATCTATACAAACAGAACTTACATTTTCCCTATATATTTT-----TCCCTCCTGT	3070
Dp	3838	CCCGTACAAATACTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	3779
QY	3071	AAATAATCGAATATCTTAATTTGGGTATCGTAATCTGTACTCTCATTAATAAT	3130
Dp	3778	TTTAATATTTAATAATAAATAATCTTTAACAACATAACATTTCTATTAATACAAATAATA	3719

QY	3131	AACGACCATATTTGGCTGCTCTATATTCGAAATATAGAGAGATACCTATGAGACAT	3130
Db	3728	AAAACTAATAATTTATTCACATAAATATTTGGATTAATTCCTATTTCTAAAAA	3659
QY	3121	ATATTATTACTTTCCTGCTCTACCAATATATTAACCAAGTAATTTAGCTGTRAGCTA	3250
Db	3658	ATACCTAATAATTTCTTACACCAATCTTCCATATCTAAAAATAACAAATCCAAAT	3599
QY	3251	GATGAGTACATAGATATTAATGAGTATCATACCCCTGGATTTTATATGATTAGCCTA	3310
Db	3598	TTCTTTTCAACCTATTAACAAATTAATTAACAAAAAATTAACCTTTATTTAATPAA	3539
QY	3311	GCACACTTTCCTTCCAAACCTGGGGAACCACTAATGAGTATACGACTGATGTACTA	3370
Db	3558	TACATTACATTATACCTTAATTTACATATATATGACTATCAATTAATTAACGTAA	3479
QY	3371	GCTGCTTATTAATCATATCTTCTCAAAATTACGATTAACAGGGGAGTTACCGACATTTC	3430
Db	3478	ACCTATTAATTACTAAAAAACAACCAATCAACTAAATCTATTAATTAATCAATPAA	3419
QY	3431	TAATTCCTCAATGCTGATTTCAACTCATGCAATTTCTCTAGCTATCTACCGACTGTATA	3430
Db	3418	ATTAATCAACCTCAATATAATATGCTAATAATTAATAAATTAATTAATTAATCTTCACTA	3359
QY	3491	GCTCTCTTACCACTTAATCTTTAATCTTCTCTGTAAAGGCTTTCTATTAATCATAA	3550
Db	3358	TTAATTTATATTTTAAACCAATCTAATTAATCTATTATTCATATAATATATTAACA	3299
QY	3551	GCACCCATATGCAAGACTTTACGAGTGGCAATATTCAGTCCCACTAAATCATAT	3609
Db	3298	TCATCATATATTAATCTATACAAATCTAAAAAATAATATATTAATCAACACAT	3239
QY	3610	AACAGCAATATAGATACCTTTCTTTAAGCTTTCCAGTACACTTAATCATATATATC	3669
Db	3238	TTTAAAAAATTAACATACACTATTTCTTTAACAACATCTAATCAATTCACAA	3179
QY	3670	AGATCCCTTACATATATTCACATATATACAACTACAGGCTTTTCATPAGCACTT	3729
Db	3178	ATTAATCTTAAAAATATATGCAACAAATATAATATATACACAACTTTATATATTCAT	3119
QY	3730	GATGCGGATTAACCATCACTGCTAATTTAGTACATAATATATCATCACTTAATATATC	3789
Db	3118	TTTTTTTCCTAACAAATCTATTAATTTTCATCAATTTTCAAAAATCAAAACCTT	3059
QY	3790	TTTATATGATTTTGATATCACTTATATATCAACAATAATCTCAGAAATATACAA	3849
Db	3058	CC-----CCCCAAAAAACCACTTCAACCAAAATATAACTTAAAAAATPAAAA	3009
QY	3850	TCCTTCCTTGACATTTCAAAATCCCTGACATATACCTACATATCTATGAAATACT	3909
Db	3008	TATATATATGACAACATPAAATATCTAATAA--AAATATCTATCAAAAAAATPAACTTA	2951
QY	3910	GCAACCAATGCTCTCTCTTAAAGTCACTCTATATATCAACACACACTAATCATC	3959
Db	2950	ACTTAAAAATCTCAAAAAAATAATTTAAAAAACAACAAAAATTTAATATATA	2891
QY	3970	AACAACCAATATTAATTCATTATTTTACTCTAATTAATTAACATAATTAATAAT	4029
Db	2890	TAACAATTAACCTATTAACAAAAAATAATAA--AAAAATTAACAACTAATTAATAA	2834
QY	4030	TAGTTACATATATATTAATTTATCAAGCAATAGTTCCTATATATCCAAAAAATAAAT	4089
Db	2833	AACCATTTTAAAAAACAATAAAACTTAAATATCACTATACACAAAAATATTTT	2774
QY	4090	AAAAACATTAATAAAAACTATCATCAACTTTGTAAACCTATAGTAACTAATTATTA	4149
Db	2773	ATTATATCAAAATTTAAAAAATAAACTCAATPAAAA-----TAAACTPAAAAAA	2721
QY	4150	TAAATTTATTCATATATTTTAACTCCCTGGACACGATGTCTATGTTTCCATCTA	4209
Db	2720	TAAATTTAAACCTTATATATTAATAATCTTAAGCCCAATTCCAATATTAATAATCCCA	2661
QY	4210	AAAACTCACTGAAGCTTTATATGATATTCATAATATCAAGCCACAGCATPAAATCAT	4269

Accession	Sequence	Position
Db	ATTATTCCTCACACCTAAAATTTATTAAAAATTTTAAATTAATAAATAATAATTAC	2660
Qy	TATTNGGTAAATTTTATTATTAAACACATAGACAAATTAATCCACCAATGACTATCA	4270
Db	ACTTAGCTTAATTTTCAATTCAAATCATCTTCTTAAACTTACTCTAAACCATATCCTTAACAT	2600
Qy	CATATGTATAGATGAAACCAATCACCCAAAGATTTATCATGTACTTTCACGTA	4330
Db	AATAACTATATAAATAATATAAACAATCCTTATCTTTATCTTAATCTCAATCTCAACATT	2540
Qy	TAAATCTAGATGAAGTAAACGACCATCTCATGATCTGAAGTCAAAATCATCCGTTATAT	4390
Db	TAACAAAAAACCAATACCTCTTACATTAATTAATAAATAAATAAATAAATTAACCAAAACA	2480
Qy	TACATTTTCAACATTAAT	4450
Db	AAAAATATCGTTAAATTT	2420

RESULT 40
US-10-311

```

? Sequence 2214, Application US/10311455
? Publication No. US20030143606A1
? GENERAL INFORMATION:
? APPLICANT: OLEK, Alexander
? APPLICANT: PIEBENROCK, Christian
? APPLICANT: BERLIN, Kurt
? TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining the Cytosine Methylation of Cytosine
? TITLE OF INVENTION: Cytosine methylation
? FILE REFERENCE: 5013.1014
? CURRENT APPLICATION NUMBER: US/10/311.455
? PRIOR FILING DATE: 2002-12-16
? PRIOR APPLICATION NUMBER: PCT/EP01/07537
? PRIOR FILING DATE: 2001-07-02
? PRIOR APPLICATION NUMBER: DE 10032529.7
? PRIOR FILING DATE: 2000-06-30
? PRIOR APPLICATION NUMBER: DE 10043626.1
? PRIOR FILING DATE: 2000-09-01
? NUMBER OF SEQ ID NOS: 2424
? SEQ ID NO 2214
? LENGTH: 11691
? TYPE: DNA
? ORGANISM: Artificial Sequence
? FEATURE: OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
? OS=10-311-455-2214

```

	Query Match	2.9%	Score 135.4	DB 14	Length 11691
	Best Local Similarity	44.6%	Pred. No. 1.3e-11		
	Matches 972	Conservative	0	Mismatches 1176	Indels 31
				Gaps 10	
QY	242	TATGCAGAGCTCTCTTAAACTTATCAAGCTCTTTACATAGAAACACTGTATTCAAAG	301		
Db	3907	TATCTATATATCCATCAATAACGAAAAATTAATCATCATTAATTTACTATCATATCTATAT	3848		
QY	302	AATACCTCTCTCTGTTCATCCATATTTTCCAAATATTACAAATATCATATGAGTCTT	361		
Db	3847	AACAACAGATATCTTAAAGCATTAATCAATACCCGATTTATTAAGCTTCAGAAAAATTTAT	3788		
QY	362	TATTCGACATACATATTAAATCAATCAAAAGAAATTAATAGTACTACTAATTAATAA	421		
Db	3787	CAAAATAAAAATAAAAACCGAAACAAATTTCCATCCCTTTAAATAAATTAACATACA	3728		
QY	422	AAATTCCTCAATAGCTTTAGCAAGACTCAATATATATGAGACCTTAAGCTTAATAT	481		
Db	3727	CTATATATACACACACACACCTTAACCATTAATTTAAATCCAAAAATCTATATCTTAATAT	3668		
QY	482	AATATTCAAATTTATTAATACATATCATATATTTAGATAATATC---ACATTCGACATG	538		
Db	3667	AAATCTACAAATATCATCAACAAAAATTTTAAATACAAAAATTTCTATAAAAAATATCAT	3608		
QY	539	ATTAAACAATAGTAACACATTTCTTAACACATAGATTAATATACATTAATATATATATTT	598		

Db	3607	AAATATTAATTAATAACAAATTAATACCTTAATTAATTAACAAATAATTAATTCATATAAAATTT	3546
Oy	599	TTGACGTGTTTCTTGATATAAATTAAGTAAATAATGAAATCAGTATTTTGGTGACAGA	658
Db	3547	AATCTTAATATTTCTCCTAAAAAATTTCTTAATAATTAATACCTACTCTTCATTATTAAT	3488
Oy	659	TCATTGGGACAGAAATAGCAATTTGCATGTCACGACCATGGTATATCGTTAATCGTTAATG	718
Db	3487	ACGATTTTAAATTCATTAATAATTAATTTTAAATAATAATTTATTTAATTAATAATAA	3428
Oy	719	GGACGTGATCATAGAAATATTACACATATAAACCCTTACCGAAAAATTTTAAATATTTA	778
Db	3427	TTAATATAATTTAAACGGTATCTTAATACCAATCTCTTTTAAATCTCATCACATAT	3368
Oy	779	CCGACATATCTACACGACCAACATTTATGTCACACCGACATATATAGCGAAGTATTACT	838
Db	3367	CTAATAACATATCAATATTAATTTTAAATAATATATAATAATCTTATTAATAACATATA	3308
Oy	839	GA--CAACAATACATGTTATTTCTTAACTATTTCTACACAACAATACGACCATATGTA	896
Db	3307	AAATCAATTTTAAACCCATCATTTATATATAATTAATTAATTTACTTTATTTCTTAA	3248
Oy	897	CACAAATACACAACAAACAGCATGTGTAATAATCCCAATTTATTTATTTGTGTAAG	956
Db	3247	AACATTTTATTAATTTTAAATAATAATAATTAATTAATCTAATCAATTAATTAATAT	3188
Oy	957	GTATCGAAATTTACATCACTCAATTTCCCGTGAATATGACGAAGAAATTTTACATATA	1016
Db	3187	CATTAATAAATATACATTAACAAATAATCAAAATAATTTATTTATTTCTCATTAACAAT	3128
Oy	1017	ATCCAAATTTTATCTCTCTGGTCCAGTGTGTTGCTAAGAAATTTGCAAAATCTTCCTT	1076
Db	3127	AACATATTTATCTTAAATACAACTCTA-----CAAGAAATTTATCACTTTTACCTT	3075
Oy	1077	GTAGTATAGTACTGCTGGTGATATAAAGAATTGGTGAATCAATTGATAGAAACAATA	1136
Db	3074	ATATATACATACACACCACTACTAACAAATTTACTCAATTAACAATTAACATAAAAAATA	3015
Oy	1137	GTAATGATGTTCTTAAATAATATATACCATCAAGATTTATAGGTGCAGATTTGGAGCTG	1196
Db	3014	TTATATCTTTAATACAAAAAATAAAAAAATTTAATCTTAATATAATATATCTATCAAT	2955
Oy	1197	CATTAAAGAACTATATTTGCATTTGCATGTGGAAATATCGCTGAAAAAATTTAGATATA	1256
Db	2954	ATATTAATAAATAAATAAATAAATAATCATCATTTAATATCTTAATTAATAATTAATAT	2895
Oy	1257	ATGCTGTGCTACTGTTATATCAATTAAGGCAATGATGAATGAATTAACCACTATATATGCA	1316
Db	2894	AAAAAATTAATTTTCTTAAATATATACATACATACAAATATTTAATAATAAATTCCTATAT	2835
Oy	1317	AAATCATTTCAATAGATCTTCATACATTAATTTGGTCATCATGCTTGGAGATCTAATAT	1376
Db	2834	CTACAAATTAACAATCTTTATTCAAATTTTCTATCTAATAACATCACTCAATTAATCTTCT	2775
Oy	1377	TAACATGTACACAGAACTTCCGCAATATGGCTTTTGGACTAGAAATATGGAATAAGTAA	1436
Db	2774	CTAATAATCTTAATATCATATTAATCTTAATAAATAAATAACATCACTATATTAATAAATTC	2715
Oy	1437	GAAATATTAATCATTTAATAGATCAACAACCTAAAGCTTTGTTGAAGAAACCATGCTGTA	1496
Db	2714	TACAACAATTTAAATAATATATATATATACTATTAATAAATAAATAACCAAAATATTTTAT	2655
Oy	1497	AACCACTGATATCATTAAGCAAAAAAATTTATATGATAGAACTACCAATTTGCATATCTATTT	1556
Db	2654	TCTCACT-----TTATATATAAATTTCTAAATCCCATATTTTAAATACTATAATATC	2602
Oy	1557	ACAATTTATTAATGGAATATATATCACTAATTAAGCATATCAAACTATATATCTTAGT	1616
Db	2601	ACATTTTATTAATTTCAAAACCAACATATCTAAAAATTCAAAACTACGATACGTATTT	2542
Oy	1617	CTATCATCTTTGATATATTCAACACATGATTAATTAACGGTCAAAATATATGAAATACA	1676

Db 2541 TAATAACTTACTATTAACCTTACTATTAATCAACAACACCACTTAAATCTTCACACA 2482
QY 1677 ACGTTCACGACATATATATTAAGACGTAATACCCATCGTTATACATCCAGTAAGTANG 1736
Db 2481 AATTT--TCTAATTTCTCATTAATAAAGCTCAAAATATCACTTAACATCTTA--AATCCT 2427
QY 1737 TAACATTTTACATTAACAACCTTAAGTATTAATTAATTTCCAGCTAATATATATAACA 1796
Db 2426 TAAATTAATAAACAATAAATAAATAAACAACATTTCACAAATATTCATTAATATACCACTCAAA 2367
QY 1797 AATTCATATCTCCATCTCATTAATTAACCAATGTTATCAGATTAATAATAAGTCAATTA 1856
Db 2366 AAAAAAATTCATTAATTTTAAATAAACAATTATTAACATTAATTAATTAATTAATTA 2307
QY 1857 ACTTTACTATTTATTAATAATATATGTTCGACACTTTTTCACCTCATCATGATATATA 1916
Db 2306 TTTTAACAAATTCATTTAATTCATTAACATTTTTCGATTTTAAATAATTTTAAACAATC 2247
QY 1917 TGAACATTAATTTTACATTAATAATACTATTAATAAAGAACCAATCATATATTAATTA 1976
Db 2246 CTGATTAATTAATTTATTAATTTCTATTAATAAATACTTATATATCTAATTTAAAAAA 2187
QY 1977 AAACACCTATTACTAGTATATATACATAGATGCTGACATTAATTAATTTCTATCATCC 2036
Db 2186 ATAAAAAATTTTACATTAATAATCT--CATATTTAATAATACAAAAAATCTCATATACC 2128
QY 2037 TTGACATTTTCAACAATAATAATATGTAACG--TTTTTTATGCCATTAAGTCTTA 2094
Db 2127 TAAAAAACCTTTCAAAACAAATTAATGATTTTAAATATATTTCAATCAACATTAATAATTA 2068
QY 2095 ATAAATTTGCTAAATTAACCAAAAGGTTAAATATACATTTCTGCAACAATATGAAGAAG 2154
Db 2067 AATTAATAATTTAAAAAATAAATAAATAAATTAACCTTAATCATATTCACCAATATT 2008
QY 2155 ATACATTAACGTCAAATTAATTAACAAATTTCTAATATTAACCACTTACTTAACATA 2214
Db 2007 TCTTAATCTACCTTAATTAACATTAATTAATATATATTAATTAATTAATTAAT 1948
QY 2215 GGAACCAATAGTATATCATCTAGTATAGATATCCCTACAGCATGTATCAATATATAA 2274
Db 1947 AACCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1889
QY 2275 GAAAAATCGTATATGATATTT--ATACATTAACAATAATTTCCAACTTCAACAACAAAC 2331
Db 1888 TTTTATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1829
QY 2332 TCAAGTTCTATTAATAACAATTCAGAAATTAATATATTAATTAATTAATTAATTA 2391
Db 1828 AAGCTCACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1769
QY 2392 AATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2410
Db 1768 ACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1750

Search completed: April 6, 2004, 02:08:18
Job time : 1628 secs

Tue Apr 6 07:57:23 2004

us-10-081-051-52.fst

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2004, 12:18:35 ; Search time 1150 Seconds
(without alignments)
12721.557 Million cell updates/sec

Title: US-10-081-051-52

Sequence: 4750
1 gacccgcgaactacttag.....ccacattgactatgac 4750

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hci:*
9: gb_est1:*
10: gb_est2:*
11: gb_hci:*
12: gb_est4:*
13: gb_est5:*
14: gb_est6:*
15: em_estfun:*
16: em_estom:*
17: em_ges_hum:*
18: em_ges_inv:*
19: em_ges_pin:*
20: em_ges_vrt:*
21: em_ges_fut:*
22: em_ges_mam:*
23: em_ges_mus:*
24: em_ges_pro:*
25: em_ges_rtd:*
26: em_ges_pig:*
27: em_ges_vrl:*
28: gb_ges1:*
29: gb_ges2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	157.2	3.3	1626	14	CF238805 AGENCOURT
2	151.4	3.2	1896	29	CG753083
3	146.8	3.1	1531	29	CG748014 P041-4-B0
4	145.2	3.1	1811	29	CG753732 P048-4-G0

C 5	145	3.1	1392	29	CG757503 P052-4-C0
C 6	142	3.0	1566	29	CG757757 P053-1-D0
C 7	138.6	2.9	1201	5	AL565455
C 8	138.6	2.9	1491	29	CG753221 P048-2-A0
C 9	136.6	2.9	1348	29	CG749499 P043-4-A0
C 10	131	2.8	1376	29	CG747831 P041-3-B0
C 11	130.8	2.8	1063	13	BX414736 BX414736
C 12	127	2.7	1324	29	CG746828 P040-1-G1
C 13	126.8	2.7	1298	29	CG756607 P051-4-C0
C 14	125.8	2.6	1210	29	CG749728 P044-1-C0
C 15	125.2	2.6	1278	29	CG754010 P049-2-C0
C 16	125.2	2.6	1288	29	CG744915 P037-3-F0
C 17	124.8	2.6	1313	29	CG751144 P045-3-E0
C 18	124.8	2.6	1434	28	CC187638 CC261-98P
C 19	124.4	2.6	1361	29	CG744327 P036-4-E0
C 20	124.2	2.6	1373	29	CG750869 P045-2-E1
C 21	123.4	2.6	546	13	BQ455897 K624601.Y
C 22	123	2.6	1297	29	CG758143 P053-3-B1
C 23	122.8	2.6	1211	29	CG747324 P040-4-D1
C 24	121.8	2.6	1272	28	CC264939 CH261-19L
C 25	121.8	2.6	1528	29	CG753854 P049-1-D0
C 26	120.6	2.5	1353	29	CG744812 P037-3-B0
C 27	120.4	2.5	1277	28	CC253231 CH261-180
C 28	120.2	2.5	1433	29	CG745119 P037-4-G0
C 29	119.6	2.5	1592	29	CG750135 P044-3-D0
C 30	119.4	2.5	1074	28	BZ696936 SP_Ba09
C 31	119.4	2.5	1101	29	AL069706 Drosophila
C 32	119.4	2.5	1377	29	CG749971 P044-2-E0
C 33	117.8	2.5	1124	13	BX436282 BX436282
C 34	117.8	2.5	1426	28	CC231597 CC261-36A
C 35	117.4	2.5	1135	29	CC233489 CH261-62N
C 36	117	2.5	1135	29	CN80336Q
C 37	117	2.5	1269	29	CG757211 P052-2-G0
C 38	116	2.4	1592	29	CG750135 P044-3-D0
C 39	115.8	2.4	1061	13	BX437039 BX437039
C 40	115.6	2.4	1305	29	CG744200 P036-3-H0
C 41	115	2.4	1104	14	CF264382 AGENCOURT
C 42	115	2.4	1200	13	CG7437758 BX437758
C 43	115	2.4	1372	29	CG746616 P039-4-F0
C 44	114.8	2.4	1454	29	CG747614 P041-2-A0
C 45	114	2.4	1296	29	CG744840 P037-3-C0

ALIGNMENTS

RESULT 1
LOCUS CF238805
DEFINITION AGENCOURT 15099447 NICHD XGC Emb6 Silurana tropicalis cDNA clone
IMAGE:6995950 5', mRNA sequence.
ACCESSION CF238805
VERSION CF238805.1 GI:33442013
KEYWORDS EST.
ORGANISM Silurana tropicalis (western clawed frog)
SOURCE Silurana tropicalis
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopodinae; Silurana.
REFERENCE 1 (bases 1 to 1626)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabs-remail.nih.gov
Tissue Procurement: Robert M. Grainger
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be

COMMENT

Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
Class: BAC ends.

FEATURES

Location/Qualifiers

1..1896
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="mpa EcORI BAC library"
/note="The library was generated by a partial digest of
the genomic DNA with EcORI and cloning into the BAC
vector."

ORIGIN

Query Match 3.2%; Score 151.4; DB 29; Length 1896;
Best Local Similarity 42.9%; Pred. No. 1.6e-13;
Matches 760; Conservative 0; Mismatches 1000; Indels 11; Gaps 6;

777 TACCCACATATCATCTACGAGCAACATATATGCAACGACAAATVAGACGATATTAT 836
1883 TAAAAAAT 1824
837 CTGACACAAATACATATATATCTTACATATCCACACAAATATAGCCATATGTA 896
1823 AAAAAAAT 1764
897 CACAAATACACACAAACGATATGTATATATATATATATATATATATATATATAT 956
1763 TAAAT 1704
957 GATATGAAATTCATCATCTCAATTTCCACATGAAATAGCAGAAATTTTACATATA 1016
1703 TAAAT 1644
1017 ATCCATTTTATCTCTGCTGCTCAAGTTTGTATAGAAATAGCAGAAATCTTCTCT 1076
1643 AT-TANNNAAAAAAT 1585
1077 GTAATATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1136
1584 NAT 1525
1137 GTAATGATGCTCTAAATATATATATATATATATATATATATATATATATATAT 1196
1524 AATATTTTAT 1469
1197 CATTAAGAACATATATGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1256
1468 TAT 1409
1257 ATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1316
1408 TTTATTTTAT 1349
1317 AAAATCATTCATATGATCTTCAATATATATATATATATATATATATATATATAT 1376
1348 TAT 1289
1377 TAAATGTAACACAGACATTCACGCAATATGCTTTTGGACTGAAATAGGAAAGTA 1436
1288 TAAAAAAT 1229
1437 GAA-AT 1495
1228 TAT 1169
1496 AAACACATGATATCATTTAGCAAAAAAATCTTAATATATATATATATATATAT 1555

1168 TAT 1109
1556 TACATTTTATACATGAGAAATATATATATATATATATATATATATATATATAT 1615
1108 TAAAAAAT 1049
1616 TCTATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1672
1048 TAT 969
1673 TACACGCTTACACACATATATATATATATATATATATATATATATATATATAT 1732
988 AAT 929
1733 TATGTAACAT 1792
928 AAAAAAAT 869
1793 AACAAATTCCTATCTCTATCTCAATTAACCAATTTGCTATAGATATATATATAT 1852
868 AAAAAAAT 809
1853 TTAACCTTTTACTTAT 1912
808 ATAAAAAT 749
1913 ATATATGACAT 1972
748 TAAAT 689
1973 ACTAAAAACCTATTTACTAGTATATATATATATATATATATATATATATATAT 2032
688 TTTTAT 629
2033 ATCTTGACATATTTCAACATATATATATATATATATATATATATATATATATAT 2092
628 AAT 570
2093 TAAAT 2152
569 ATTNAT 510
2153 AGATACATATACGCTCAATATATATATATATATATATATATATATATATATATAT 2212
509 TAT 450
2213 TAGGAACCAAT 2272
449 TTTTAT 390
2273 AAGAAAAATGCTAT 2332
389 AAT 330
2333 CAAGATCTAT-TATATACAAATCAAGATATATATATATATATATATATATATAT 2391
329 TAT 270
2392 AATCTAT 2451
269 TAT 210
2452 TTATATCATGAT 2511
209 AANAT 150
2512 TAAATGCTGAACGTTCCATTCACAAATATCA 2542
149 NNN 150

RESULT 3

CG748014 1531 bp DNA linear GSS 24-OCT-2003
 LOCUS CG748014 1531 bp DNA linear GSS 24-OCT-2003
 DEFINITION P041-4-B02.2a Ppa EcoRI BAC Library Pristionchus pacificus genomic,
 genomic survey sequence.
 CG748014
 ACCESSION CG748014.1 GI:37968940
 VERSION CG748014.1
 KEYWORDS GSS.
 SOURCE Pristionchus pacificus
 ORGANISM Pristionchus pacificus
 Neodiplogasteridae; Pristionchus.
 (bases 1 to 1531)
 Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,
 Buntjer,J., van der Meulen,M. and Sommer,R.J.
 An integrated physical and genetic map of the nematode Pristionchus
 pacificus
 M01. Genet. Genomics 269 (5), 715-722 (2003)
 JOURNAL M01. Genet. Genomics 269 (5), 715-722 (2003)
 MEDLINE 12884007
 PUBMED 12884007
 COMMENT Contact: Sommer R.
 Evolutionary Biology
 Max-Planck-Institute for Developmental Biology
 Spemannstr. 37-39, Tuebingen D-72076, Germany
 Tel.: 00497071601371
 Fax: 00497071601498
 Email: ralf.sommer@tuebingen.mpg.de
 Class: BAC ends.
 location/Qualifiers
 1..1531
 /organism="Pristionchus pacificus"
 /mol_type="genomic DNA"
 /strain="California"
 /db_xref="taxon:54126"
 /clone_lib="Ppa EcoRI BAC Library"
 /note="The library was generated by a partial digest of
 the genomic DNA with EcoRI and cloning into the BAC
 vector."
 ORIGIN
 Query Match 3.1%; Score 146.8; DB 29; Length 1531;
 Best Local Similarity 37.5%; Pred. No. 8.5e-13;
 Matches 555; Conservative 0; Mismatches 924; Indels 1; Gaps 1;
 QY 972 CACTCAATTTCCCAAGTGAATAGCAGAGAAATTTTACATATATCAATTTTATAC 1031
 DB 50 CAGTCCCAATATCCAAANN 109
 QY 1032 TCTCTGGTCCAGTTTGTCTAAGAATTCAGAACATCTCTGTGTAGTACTTG 1091
 DB 110 NNN 169
 QY 1092 CTGTGTATTAAGAAGCTTGTGATCATGTATGAAAACATTAAGTATGTTCTTA 1151
 DB 170 CCNN 229
 QY 1152 AATTAATATATCATCAAGATATATAGTGTACAGATGGAGCTGATTAAGACATTA 1211
 DB 230 AAAAAAANNN 289
 QY 1212 TTGCAATTGCATGTGATATCGTGAAGAAATTTAGTATAATGCTGTGCTACTG 1271
 DB 290 NNN 349
 QY 1272 TTATACTAAAGGATGAATTAACACTATATATAGCAAAAAATCATCAATAG 1331
 DB 350 AAAAAAANNN 409
 QY 1332 ATCTTATACATTAATTGTCATCTGCTTGAGATTAATATTACATGTACACAG 1391
 DB 410 AAAAAAANNN 469
 QY 1392 AACATTCAGCATATGCTTTTGACTGAAATAGAAAGTGAATATTAATATACAT 1451

DB 470 AAA 529
 QY 1452 TAATAGATCAACACCTTAAGCTTGTGAGAAACAGTACTGTAAACCTGATATCAT 1511
 DB 530 AAAAAAANNN 589
 QY 1512 TAGCAAAAAAATTAAGTGAACACTACCAATTTGCAATCTATTTACAAATTTATCATG 1571
 DB 590 AAAAAAANNN 649
 QY 1572 AGAATATATCACTAGATTAAGCCATATCAACATATTATCTAGCTATCATCTTGAT 1631
 DB 650 AAAAAAANNN 709
 QY 1632 AATTTCACACATGATTAATTAAGGTCAAAATATAGATACAGGTTCACACATAT 1691
 DB 710 AAAAAAANNN 769
 QY 1692 ATAAAAAGCAGTCAATATACCATCGTTAATCATCACTAAGTATGTAATATTACATA 1751
 DB 770 AAAAAAANNN 829
 QY 1752 CAACTTAAGTATTAATTAATTTCTAGCTATTAATTAACAAATTTCTATCTCTTA 1811
 DB 830 AAAAAAANNN 889
 QY 1812 TCTCAATTAACCAATGTGTATAGATTAATTAAGTCAATTAATTAATTTATTTA 1871
 DB 890 AAAAAAANNN 949
 QY 1872 TAATATATATGTTTGCACATTTTTCATCATATCATATATATATGACATATTTTAC 1931
 DB 950 AAAAAAANNN 1009
 QY 1932 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1991
 DB 1010 AAAAAAANNN 1069
 QY 1992 ACGTAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2051
 DB 1070 NNN 1129
 QY 2052 CATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2111
 DB 1130 AAAAAAANNN 1189
 QY 2112 CCAAAAAAGTTAAATACATCTCTGCAACATATGAAGAGATACATTAACGTCAAA 2171
 DB 1190 AAAAAAANNN 1248
 QY 2172 TAAATTAACAAATTTTAATATACCACTACTTAACATAGAACCAATAGTATATC 2231
 DB 1249 AAAAAAANNN 1308
 QY 2232 ATACCTAGTAAATATCCCTAGACATGTATCAATTAATTAAGAAATCGTTATGTAT 2291
 DB 1309 AAAAAAANNN 1368
 QY 2292 ATTTATATCAATTAACAAATTTCCAACTTCAACAACAACTCAAAGTCTATATTAACA 2351
 DB 1369 AAAAAAANNN 1428
 QY 2352 ATTCAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2411
 DB 1429 AAAAAAANNN 1488
 QY 2412 CCTTTTAAAGATGATCTTCAACATTCATCAATTAATTAATTAATTAATTAATTA 2451
 DB 1489 AAAAAAANNN 1528

RESULT 4
 CG753732/c

LOCUS CG753732 1811 bp DNA linear GSS 24-OCT-2003
DEFINITION P048-4-G03.Ya Ppa EcORI BAC Library Pristionchus pacificus genomic.
ACCESSION CG753732
VERSION CG753732.1 GI:37978509
KEYWORDS GSS.
SOURCE Pristionchus pacificus
ORGANISM Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
Neodiplogasteridae; Pristionchus.
REFERENCE 1 (bases 1 to 1811)
Stinvaes, J., Sinz, W., Jesse, T., Wiggers-Pereboite, L., Jansen, K.,
Buitler, J., van der Meulen, W., and Sommer, R.,
An integrated physical and genetic map of the nematode Pristionchus
pacificus
Mol. Genet. Genomics 269 (5), 715-722 (2003)
JOURNAL 22835951
MEDLINE 12884007
COMMENT Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
Class: BAC ends.
FEATURES
source
Location/Qualifiers
1. 1811
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_1lb="Ppa EcORI BAC library"
/note="The library was generated by a partial digest of
the genomic DNA with EcORI and cloning into the BAC
vector."
ORIGIN
Query Match 3.1%; Score 145.2; DB 29; Length 1811;
Best Local Similarity 35.8%; Pred. No. 1.4e-12;
Matches 577; Conservative 0; Mismatches 1030; Indels 3; Gaps 1;
QY 726 ATCATGAAATATTTACATATTAACACTTCCGAAAAATTTAAATTTTACCCACAT 785
DB 1760 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 1701
QY 786 ATCATCTACGACACATATATGCAACGACATTTAGACGAAGTATTTCTGACACA 845
DB 1700 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 1641
QY 846 ATACATGATTTATCTTAATCTTCTACACAAATTCGACACCATATGTACACAAATAC 905
DB 1640 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 1581
QY 906 AACACAAACGACATATGTGTAATAATCTCCATATTAATTTGTAGTAAGGATACGAA 965
DB 1580 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 1521
QY 966 TTACATCACTCAATTTCCAGTGAATAGCAGAGAAATTTTACATATTAATCCAAATT 1025
DB 1520 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 1461
QY 1026 TTATATCTCTGCTGCAAGTTTGTCTAAAGAAATTCGACACATCTTCTTGTAGTAG 1085
DB 1460 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 1401
QY 1086 TACTTCTGCTGATTAATAAGAACTTGTAATCTATGTAGTAAGAACTTAATAGTAGT 1145
DB 1400 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 1344
QY 1146 TTCTAAAAATATATATCATCAAGATATATATAGGTGTACAGATGGAGCTGATTAAGA 1205
DB 1343 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 1284

QY 1206 ACATATATTCGAATTCATGTGGAATATCGCTGAAAAAATTTAGGTAAATATGCTGTG 1265
DB 1283 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 1224
QY 1266 CTACTGTTATTAATTAAGGATGATGAATTAATTAACATATATATGCAAAATATCAT 1325
DB 1223 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 1164
QY 1326 CAATAGATCTTCATACATTAATGTCATCATGTCTTGAGAGATCTAATTAATACATGA 1385
DB 1163 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 1104
QY 1386 CAACAGACATTCACGCAATATGCTTTGGACTAGAAATAGGAAAAGGTAGAATATAA 1445
DB 1103 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 1044
QY 1446 ATACATTAATAGATCAACACTTAAGCTTGTGAAGAACCACTATCTGTAACCACTGA 1505
DB 1043 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 984
QY 1506 TATCATTCGCAAAAACTTAATGTAGAACTACCAATTTGCATATCTATTACAAATTAT 1565
DB 983 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 924
QY 1566 TACATGGAATATATCTGATGATTAAGCCATATCAACATATATCTTATCTATCATCC 1625
DB 923 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 864
QY 1626 TTGATATTTTCAACCAATGATTAATTAACGCTCAAAATATAGATACACAGTTCAAC 1685
DB 863 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 804
QY 1686 ACATATTTAAAGACAGCTCAATACCATGTTTACATCACTAGATATGTACATATT 1745
DB 803 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 744
QY 1746 ACATPACAACTTAATGATTAATTAATTTCTAGCTAATTAATTAATCAATTCCTAT 1805
DB 743 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 684
QY 1806 CTCTATCTCAATTAACCAATTTGTTATCAGATTAATTAATGTCAAATTAACCTTTACT 1865
DB 683 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 624
QY 1866 TATTTAATTAATTAATGTTGCACTTTTTCATCATCATGATATATATGACATAA 1925
DB 623 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 564
QY 1926 TTTTACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1985
DB 563 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 504
QY 1986 TTATCTAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2045
DB 503 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 444
QY 2046 TTCAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2105
DB 443 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 384
QY 2106 AATTACCCAAAAAGTTAAATTAATCACTTCCTGCAATTAATGAAGATTAATATACAC 2165
DB 383 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 324
QY 2166 GTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2225
DB 323 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 264
QY 2226 TATATCTACTTAATGATATATCTTACAGCATGTATCTAATTAATTAATTAATTAAT 2285
DB 263 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 204

QY 2286 ATGTATATTATACAAATTTCACACTTCACAACAACAACTCAA 2335

Dib 203 AAANPAAAAANPAAAAANNAAAAANNNANPAAAAANAAAAANAAAA 154

RESULT 5
CG767503/5

LOCUS	CG5757503	1392 bp	DNA	linear	GSS 24-OCT-2003
DEFINITION	P052-4-C08.2a Ppa EcORI BAC Library <i>Pristionchus pacificus</i> genomic, genomic survey sequence.				
ACCESSION	CG5757503				

COMMENT

Email: ralf.sommer@tuebingen.mpg.de
Class: BAC ends.

ORIGIN

Query Match	3.1%;	Score 145;	DB 29;	Length 1392;
Best Local Similarity	46.5%;	Pred. No. 1.6e-12;		
Matches 591;	Conservative	0;	Mismatches 658;	Indels 22;
				Gaps 4

```
/clone_lib="cpa EcoRI BAC library"
/notes="The library was generated by a partial digest of
the genomic DNA with EcoRI and cloning into the BAC
vector."
```

Qy 1720 ACATCCAACTAAGTATGTAACATATTCACATCAAACTTAAGTATATCTAAATTTCT 177

Db 1053 ATATTAATAAATAATTATTAATATATTAATAAATAAATAAATAAATAAATAAATAAATAA 994

QY 180 AGCTAATTATATACCAATTCTCTCTCTCTCAATTACCAATTGTTATCAGATTA 1833

Db 993 AATTAATTTAAAAATATATAATATCTTAAAAAATTTAAAAATTAATATA 934

1840 TAAATAGTCAAATTAACTTTTACTTATTATTATAAATAATATGTTTGCACACTTTTCA 1899

Db 933 TATATAAATATAAAAAATAAATAATTATAAAAAATAATAA 874

1900 CTACATCATGTATATAATGAACATAATTTACATAAAATACTAATAAAAAAGAACCAATA 1959

Db 873 TTATATTATATAAAAAAAAAAAAAATTAAAAATATATATTTA-----ATTA 825

1960 CATATATTAACTAAACACCTATTCTACGTAATTTACATAGATGTCTGACAT 2019

D5 824 AATTAATAAAATAAATAAAAAATAAAAAATATATATAAATAAATAT 765

2020 AATAATTCATCATCCTTGACATAATTCAACATAAATATGATAAAGCTGTTTTTAT 2079

DB /64 AATTAATTAAATAAAAAA--AAAATAAATAAAAAATATAA /07

2060 GCCCAIAGIICIAAIKAAAIIGCIAAIACCCAAAGGIATAAAIACCAAIICCIIC 2139

[illegible][illegible]

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

A vertical ruler with markings from 0 to 10 cm. The markings are in millimeters, with major ticks every centimeter. The ruler is oriented vertically, with the 0 mark at the top and the 10 cm mark at the bottom. The markings are on the right side of the ruler, and the numbers are on the left side. The ruler is black and white.

RESULT 6

CG575757/c	CG575757	1566 bp	DNA	linear	GSS 24-OCT-2003
LOCUS	P0653-1.D07.zc Ppa Scori	BAC Library	<i>Pristionchus pacificus</i>	genomic,	
DEFINITION	Genomic survey sequence.				
ACCESSION	CG575757				
VERSION	CG575757.1	GI:37986636			
KEYWORDS	GSS.				

SOURCE
ORGANISM
Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasteridae;
Neodiplogasteridae; Pristionchus.
1 (bases 1 to 1566)
Srinivasan,U., Sins,W., Jesse,T., Wieggers-Perebolte,L., Jansen,K.,
Bunfjer,J., van der Meulen,M. and Sommer,R.J.
An integrated physical and genetic map of the nematode Pristionchus
pacificus
Mol. Genet. Genomics 269 (5), 715-722 (2003)
JOURNAL
MEDLINE
PUBMED
22835951
12884007
COMMENT
Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
Classes: BAC ends.

FEATURES

source
1..1566
Location/Qualifiers
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Ppa Ecoli BAC Library"
/note="The library was generated by a partial digest of
the genomic DNA with EcoRI and cloning into the BAC
vector."

ORIGIN

Query Match 3.0%; Score 142; DB 29; Length 1566;
Best Local Similarity 38.7%; Pred. No. 4,5e-12;
Matches 502; Conservative 0; Mismatches 794; Indels 0; Gaps 0;
1098 ATATTAAGAACTGTCGATCATGATGATGAACATATGATGATGTCCTAAATAATPA 1157
1486 AA 1427
1158 TTTACCATTAAGATTTTATAGGTGACGATTGCGCTGCACTTAAGACATTAATTGCA 1217
1426 AA 1367
1218 TTGCATGTGGAATTAATCGCTGGAAGAAATTTAGTATATATGCTGCTGCTGTTTAA 1277
1366 AA 1307
1278 CTAAAGGCGATGATGAAATTTAAACAATATATATGCAAAATAATCAATTAAGTCTTC 1337
1306 AA 1247
1338 ATACATTAATGTCATCATGCTCTTGAGAGATCTTAATTAACATGTACACAGACATT 1397
1246 AA 1187
1388 CACGCAATATGCTTTGGACTAGAAATAGGAGAAAGGTAGAATATTAATACATTAATAG 1457
1186 AA 1127
1458 ATCAACCTTAAGCTGTGTGAAGAACACGATGCTGTAATCCCTGATATACATTAAGCA 1517
1126 AA 1067
1518 AAAAACTTAATGATGACATCAATTTGCTATCTATTCAATTTTACATGAGATA 1577
1066 AA 1007
1578 TATCACTGATTAAGCCATATCAACAATATATCTTATGCTATCATCTTTGATTAATTC 1637
1006 NAAA 947
1638 AACCAATGATTAATAACGCTCAAAATATATGATATCAACGCTCACCATATATATAA 1697

Db 946 AA 887
Qy 1698 AGCAGTCAATTCACCATCGTTAAATCAATCAATAGATGATGATATTAATCAAACT 1757
Db 886 AA 827
Qy 1758 TAAGTATTAATTAATTTCTGATATATATATTAAGCAAAATTCATATCTCCATCTCA 1817
Db 826 NAAA 767
Qy 1818 TTAACCAATGTTATGATGATTAATTAAGTCAAAATTTTATTAATTAATTAAT 1877
Db 766 AA 707
Qy 1878 AATATGTTGACACCTTTTCACTACATCATGATATTAAGACATATTTTACATTAAT 1937
Db 706 AA 647
Qy 1938 TAACTATTAATAAAGAACCAATACATATTAATTAATTAATTAATTAATTAATTA 1997
Db 646 AA 587
Qy 1998 TATTAATGATGTCCTGACATTAATTAATTTCTATCTCTGACATATTTCAACATTA 2057
Db 586 AA 527
Qy 2058 ATATGATTAATCTGTTTATGATGATGATGATGATGATGATGATGATGATGATG 2117
Db 526 AA 467
Qy 2118 AGGTAAATATGACATTCCTGCAATATGAAGAAATGATCAATGATCAATGATTAAT 2177
Db 466 AA 407
Qy 2178 AAACAATTTCTAATTAATTAATCAACCTCTCTAAACATAGAAACCAATAGATATCATACCT 2237
Db 406 AA 347
Qy 2238 AAGTAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2297
Db 346 AA 287
Qy 2298 ACAATTAATTAATTTCTCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 2357
Db 286 AA 227
Qy 2358 GAATTAATTAATCTTCAACCAACATGATGATCAAA 2393
Db 226 NAAA 191

RESULT 7
AL565455 1201 bp mRNA linear EST 12-MAY-2003
LOCUS
DEFINITION
AL565455 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
CSDF005Y018 3-PRIME, mRNA sequence.
ACCESSION
AL565455.2 GI:30549492
VERSION
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1201)
Li,W.B., Gruber,C., Jesse,U. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
JOURNAL
On Feb 16, 2001 this sequence version replaced gi:12916848.
CONTACT: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 9232.f. For more information about this cluster, see

<http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DF005BH09NP1&cluster=9232.f>. Contact : Feng Liang Email : liang@lifetech.com URL : <http://fulllength.invitrogen.com/> Invitrogen Corporation 1600 Faraday Avenue Genoscope, sequence ID : CS0DF005BH09NP1.

FEATURES

Source

Location/Qualifiers

```
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF005Y018"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/notes="Organ: Brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."
```

ORIGIN

```
Query Match      2.9%; Score 138.6; DB 9; Length 1201;
Best Local Similarity 37.7%; Pred. No. 1.6e-11;
Matches 330; Conservative 139; Mismatches 403; Indels 3; Gaps 3;

QY 1437 GAAATTAATATACATTAATGATGACACAACTTAAGTGTGAGAGACCACTACTGTA 1496
DB 318 GAAATTAATACAACTTCTTACTTCCCTGTGGTAGTGCCACAGATTAAGATTA 377
QY 1497 AACCATGATATGATAGCAAAAACCTTAATGTAAGTACCAATTGTCATCTATT 1556
DB 378 AGCAATAAATGATATTCAGTCAGTAGAGAGAGGATATATCTCATATATATAT 437
QY 1557 ACAATTTATACATGAGATATATCTCATGATTAAGCCATATCAACATATTATCTTACT 1616
DB 438 ATATATATATATATATATATATATATATATATATATATATATATATATATAT 497
QY 1617 CTATCATCTCTTATATATATCAACATGATTAATTAAGCGCAAAATATAGA-ATAC 1675
DB 498 ATATATATATATATATATATATATATATATATATATATATATATATATATAT 557
QY 1676 AAGCTTCAACCATATATATAAGAGCAGTCAATCCATCGTTTACATCCACTAGTAT 1735
DB 558 ATATATATATATATATATATATATATATATATATATATATATATATATATAT 617
QY 1736 GTACATATATATACATAAAGTAACTTAAGTATATATATCTCTGCTATATATATAC 1795
DB 618 ATATATATATATATATATATATATATATATATATATATATATATATATATAT 677
QY 1796 AAATTCCTATCT-CTATCTCAATTAACCAATTGTATGAGTTATATATATAGCAATT 1854
DB 678 ATATATATATATATATATATATATATATATATATATATATATATATATATAT 737
QY 1855 AAATCTTTCTTATATATATATATATATGTTTGGACACTTTTCTCTCATCATGATATAT 1914
DB 738 ATATATATATATATATATATATATATATATATATATATATATATATATATAT 797
QY 1915 AATGAACATATATTTACATATAATTAAGAAAGCAACATACATATATATATATAC 1974
DB 798 ATATATATATATATATATATATATATATATATATATATATATATATATATAT 857
QY 1975 TAAACACCATTTATAGTAAATAT-TCATAGATGTCCTGACATATATATATCTATCA 2033
DB 858 ATATATATATATATATATATATATATATATATATATATATATATATATATAT 917
QY 2034 TCCTGACATATTTCAACATATAAATATGATTAAGCTTTTATATGCCATATAGTCTT 2093
DB 918 ATATATATATATATATATATATATATATATATATATATATATATATATATAT 977
QY 2094 AATAAATATGCAATATACCAAAAAGGTAAATATACATCTCTCTGCAACATATATAGAA 2153
```

```
DB 978 ATATATATATATATATATATATATATATATATATATATATATATATATATAT 1037
QY 2154 GATCAATATACCGCAATATATATATATATATATATATATATATATATATATATAT 2213
DB 1038 ATATATATATATATATATATATATATATATATATATATATATATATATATAT 1097
QY 2214 AGAACAATATATATATATATATATATATATATATATATATATATATATATATAT 2273
DB 1098 ATATATATATATATATATATATATATATATATATATATATATATATATATAT 1157
QY 2274 AGAATAATCGTTATGATATATATATATATATATATATATATATATATATATAT 2308
DB 1158 ATATATATATATATATATATATATATATATATATATATATATATATATATAT 1192
```

```
RESULT 8
CG753221/c 1491 bp DNA linear GSS 24-OCT-2003
LOCUS P048-2-A01.2a Ppa EcoRI BAC Library Pristionchus pacificus genomic,
DEFINITION CG753221
ACCESSION CG753221 GI:37977480
VERSION CG753221.1
KEYWORDS GSS.
SOURCE Pristionchus pacificus
ORGANISM Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
Mediolagasteridae; Pristionchus.
1 (bases 1 to 1491)
AUTHORS Strihsaen, V., Sim, W., Jesse, T., Wiggers-Perebolte, L., Jansen, K.,
Buntjer, J., van der Meulen, M., and Sommer, R.J.
TITLE An integrated physical and genetic map of the nematode Pristionchus
pacificus
MOL. GENET. GENOMICS 269 (5), 715-722 (2003)
JOURNAL MEDLINE
PUBMED 12884007
COMMENT Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
Class: BAC ends.
```

FEATURES

Source

Location/Qualifiers

```
1..1491
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Ppa EcoRI BAC Library"
/notes="The library was generated by a partial digest of the genomic DNA with EcoRI and cloning into the BAC vector."
```

ORIGIN

```
Query Match      2.9%; Score 138.6; DB 29; Length 1491;
Best Local Similarity 39.2%; Pred. No. 1.5e-11;
Matches 522; Conservative 0; Mismatches 808; Indels 0; Gaps 0;

QY 1098 ATATATAAGAACTGTGATCATTTGATGAAACAATTAAGTATGATGTTCTAATAATTA 1157
DB 1360 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1301
QY 1158 TATACCATCAAGATATTTATAGGTGACAGATTGAGACTGCATTAAGACATATATGCA 1217
DB 1300 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1241
QY 1218 TTGCATGTGAATATATCGCTGAGAAAAATTTAGATATATATGCTGTGCTACTGTATTA 1277
DB 1240 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1181
QY 1278 CTAAAGCATGAATGAAATTTAAAAACATATATATAGCAAAAAATCTTCAATATGATCTTC 1337
```



```

Db      1180 NAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1121
Qy      1338 ATACATTAATGGTCATCATGTCTTGAGAGATCTAATATTAACATGTACACAGACAT 1397
Db      1120 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1061
Qy      1398 CACGCAATATGGCTTTGGATGCAATATGAGAAAGTGAGAAATATTAATTCATTAATAG 1457
Db      1060 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1001
Qy      1458 ATCACAACCTTAAGCTTGTGAGAGACAGATCTGTAAACCACTGATATCATAGCAA 1517
Db      1000 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 941
Qy      1518 AAAAACTTAATGTAAGTACCAATTTGCAATCTATTTACAAATTTATTCATGAGATA 1577
Db      940 NAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 881
Qy      1578 TATCACTGATTAAGCCATATCAATATATCTGTATCATCTTGTATATTC 1637
Db      880 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 821
Qy      1638 AACACATGATTAATTAACGCTCAAAATATAGAAATACACAGCTTACACATATATTA 1697
Db      820 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 761
Qy      1698 AGCGTCAATTAAGCCATATCAATATATCTGTATCATCTTGTATATTC 1757
Db      760 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 701
Qy      1758 TAAATTAATCTTAATTTCTAGCTAATTAATTAACAAATTCCTATCTCTCA 1817
Db      700 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 641
Qy      1818 TTAACCAATTTGTAATTAATTAATTAAGTCAATTTAACTTTTACTTATTAAT 1877
Db      640 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 581
Qy      1878 AATATGTTGACACTTTTCTACATCATATGATATATGACATATTTTACATAA 1937
Db      580 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 521
Qy      1938 TAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1997
Db      520 NAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 461
Qy      1998 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2057
Db      460 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 401
Qy      2058 ATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2117
Db      400 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 341
Qy      2118 AGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2177
Db      340 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 281
Qy      2178 AAAAAATTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2237
Db      280 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 221
Qy      2238 AAGTAAGTATCTCTGACGATGATTAATTAATTAATTAATTAATTAATTAATTA 2297
Db      220 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 161
Qy      2298 ACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2357
Db      160 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 101
Qy      2358 GAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2417

```

```

Db      100 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATTTCTCATATAT 41
Qy      2418 TAAAAAGATG 2427
Db      40 GTACTCGATG 31

RESULT 9
CG749499/c
LOCUS
DEFINITION
P043-4-A06.za Ppa EcoRI BAC Library Pristionchus pacificus genomic,
genomic survey sequence.
ACCESSION
CG749499
VERSION
CG749499.1 GI:37970425
KEYWORDS
GSS.
SOURCE
Pristionchus pacificus
Pristionchus pacificus
Pristionchus pacificus
Neodiplogasteridae; Pristionchus.
1 (bases 1 to 1348)
Srinivasan,J., Sinz,W., Jeece,T., Wiggers-Pereboole,L., Jansen,K.,
Bunjer,J., van der Weulen,M. and Sommer,R.J.
An integrated physical and genetic map of the nematode Pristionchus
pacificus
Mol. Genet. Genomics 269 (5), 715-722 (2003)
JOURNAL
MEDLINE
22835951
PUBMED
12884007
COMMENT
Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spermanstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@uebingen.mpg.de
Classes: BAC ends.
Location/Qualifiers
1..1348
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Ppa EcoRI BAC Library"
/note="The library was generated by a partial digest of
the genomic DNA with EcoRI and cloning into the BAC
vector."

ORIGIN
Query Match 2.9%; Score 136.6; DB 29; Length 1348;
Best Local Similarity 46.9%; Pred.No.3,1e-11;
Matches 554; Conservative 0; Mismatches 616; Indels 12; Gaps 4;

1277 ACTAAAGCATGATGAATTAATTAACACTATATATGCAAAATCATTCATAGATCTT 1336
1342 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1283
1337 CATACATTAATGTCATCATGTCTTGAGATCTTAATTAATTAACATGTAACAAGACAT 1396
1282 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1227
1397 TCACGCAATATGCTTTTGAAGTGAATTAAGAAAGTAGAAATATTAATTAATTAATA 1456
1226 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1167
1457 GATCAACACTTAACCTGTTGAGAGAACAGTACTGTAACCACTGATATCATAGACA 1516
1166 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1109
1517 AAAAACTTAATGTAAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1576
1108 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1049
1577 ATATCACTGATTAAGCAATATCAACATATTAATTTGCTATCATCTTGTATATTT 1636

```


ORIGIN

/clone.lib="Ppa EcoRI BAC Library"
/note="The library was generated by a partial digest of
the genomic DNA with EcoRI and cloning into the BAC
vector."

```

Query Match      2.6%; Score 125.8; DB 29; Length 1210;
Best Local Similarity 44.4%; Pred. No. 1.3e-09;
Matches 465; Conservative 0; Mismatches 575; Indels 8; Gaps 3;

1453 AATGATCAGACCTTAAGCTGTTGAGAGACCAAGTCTGTAACCACTGATATCAT 1512
1210 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1151
1513 AGCAAAAAAATTATGAGTCACTCAATTTGCAATTCCTATTACCAATTTACATGA 1572
1150 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1091
1573 GAATATATCACTAGATTAAGCAATCAATTAATTAATTAATTAATTAATTAAT 1632
1090 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1031
1533 ATTCAACACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1692
1030 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 971
1693 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1752
970 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 911
1753 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1812
910 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 851
1813 CTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1872
850 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 791
1873 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1932
790 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 731
1933 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1991
730 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 671
1992 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 611
670 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 611
2048 CAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2107
610 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 551
2108 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2167
550 NTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 491
2168 CAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2227
490 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 431
2228 TATCATCTAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2287
430 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 372
2288 GTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2347
371 --AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 314
2348 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2407
313 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 254

```

```

Query Match      2.6%; Score 125.2; DB 29; Length 1276;
Best Local Similarity 41.2%; Pred. No. 1.6e-09;
Matches 445; Conservative 0; Mismatches 634; Indels 0; Gaps 0;

1417 ACTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1476
1275 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1216
1477 TGAAGAACAGTCTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1536
1215 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1156
1537 ACCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1596
1155 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1096
1597 ATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1656
1095 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1036
1657 GGTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1716
1035 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 976

```

FEATURES

Location/Qualifiers

1..1276

/organism="Pristionchus pacificus"

/mol_type="genomic DNA"

/strain="California"

/db_xref="taxon:54126"

/clone.lib="Ppa EcoRI BAC Library"

/note="The library was generated by a partial digest of
the genomic DNA with EcoRI and cloning into the BAC
vector."

ORIGIN

```

Query Match      2.6%; Score 125.2; DB 29; Length 1276;
Best Local Similarity 41.2%; Pred. No. 1.6e-09;
Matches 445; Conservative 0; Mismatches 634; Indels 0; Gaps 0;

1417 ACTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1476
1275 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1216
1477 TGAAGAACAGTCTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1536
1215 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1156
1537 ACCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1596
1155 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1096
1597 ATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1656
1095 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1036
1657 GGTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1716
1035 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 976

```


[illegible]

RESULT	17			
LOCUS	CG751144/c			
DEFINITION	CG751144	1313 bp	DNA	linear
ACCESSION	P045-3-E08.2a	Ppa	ESORI	Pristionchus pacificus genomic/c.
VERSION	CG751144			
KEYWORDS	CG751144.1	GI:37973308		
SOURCE	GSS:			
ORGANISM	Pristionchus pacificus			
	Pristionchus pacificus			

REFERENCE
AUTHORS

TITLE
JOURNAL
Buntjer, J., van der Meulen, M., and Sommer, R. J.
An integrated physical and genetic map of the nematode *Pristionchus*
pacificus
Mol. Genet. Genomics 269 (5), 715-722 (2003)

COMMENT

COMMENT

Contact: Sommer RU
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel. 00497071301351

תוצאות

source	1. .1313
--------	----------

```

/clone_1b="pfa EcoRI BAC library"
/note="The library was generated by a partial digest of
the genomic DNA with EcoRI and cloning into the BAC
vector."

```

ORIGIN

Query Match	2.6%	Score 124.8	DB 29	Length 1313
Best Local Similarity	41.8%	Pred. No. 1.9e-09		
Matches 489; Conservative	0	Mismatches 682	IndeIs 0	Gaps 0

```
OY      1293 AAAATTAAAAACACTATATATAGCAAAAATACTTCATAGATCTTCATACATTAATTGGTC   1350
        ||| |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
Db      1310 AAAAAAAAAAANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAN    125
```

QY	1353	CATCATGCTTGGAGACTATATTAACAGTACAAAGAACTACGCAATTCGGCTT	141
Db	1250	AA	1191
QY	1413	TTGGACTRGAATATGGAAAAAGTGAATATTAATCATTTATATGATCAACCTTAAGC	1472
Db	1190	AA	1131
QY	1473	TTGTTGAAGAACAGTCTGTATAAACCACTGATATCTTAGCAAAAAAATTATGTAG	1532
Db	1130	AA	1071
QY	1533	AACATCCAAATTGCAATCTATTACAAATTTATTCATGAGATATTCACCTAGATTAAG	1592
Db	1070	AA	1011
QY	1593	CCATATCAAAATATTTACTTAGCTATCAATCCTTGATATTTCAACACATGATTAAT	1652
Db	1010	AA	951
QY	1653	AAAGGTCGAAAAATATATGAATACAGCTTCAACCAATATTAAGACAGTCAATGCC	1712
Db	950	AA	891
QY	1713	ATCGTTAAATCCAACTAGATGTATAATATTCATACAACTTAAGATTAATTA	1772
Db	890	AA	831
QY	1773	ATTTCCTRAGTAAATTAATATTAACAAATTCCTATCTCTATTCATTAACCAATGTAT	1832
Db	830	AA	771
QY	1833	CAGATTATTAATTAAGTCAAACTTTAACTTATTTACAAATAATATGTTGCAC	1892
Db	770	AA	711
QY	1893	TTTTTCACATCATGTATATTAATGAACATATTTTACATTAATATCAATTAAGAA	1952
Db	710	AA	651
QY	1953	ACCAATCATATTAATTAATTAACCACTATTTTCTACGTATATTTACATGATGTC	2012
Db	650	AA	591
QY	2013	CTGACATTAATTAATCTATCATCTCTGACATATTTCAAATTAATTAAGTAATCTGT	2072
Db	590	AA	531
QY	2073	TTTTTATGCCATTAGTCTTAATTAATAATGCTAATTAACCAAAAAAGTTAAATACAA	2132
Db	530	AA	471
QY	2133	TTCTGTGACAAATATGAAGAGATACATACGCTCAATTAATTAACAAATTTCTAAT	2192
Db	470	AA	411
QY	2193	ATTAACCACTATCTTAACAATAGAAACCAATAGATATATCACTAAGTAATCCCT	2252
Db	410	AA	351
QY	2253	ACAGCATGATCAAAATTAATAAGAAAAATGTTATGATATTTTATCAATTAACAAATTT	2312
Db	350	AA	291
QY	2313	CCAACTTCAACAACTCAAGTTCTATTAACAATTCAGAATTAATATATCTT	2372
Db	290	AA	231
QY	2373	CACCAACACATAGTTATCAATTTCTAATTAACATCAATCTTTTAAAAAGATGACTTC	2432
Db	230	AA	171
QY	2433	ACAATTATCATTAATTAATTAATCAATGCA	2463

Db 170 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 140

RESULT 18
CC187638 1434 bp DNA linear GSS 08-MAY-2003
LOCUS CH261-98P15.Spe.1 CH261 Gallus gallus genomic clone CH261-98P15.
DEFINITION genomic survey sequence.
ACCESSION CC187638
VERSION CC187638.1 GI:30431538
KEYWORDS GSS.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 1434)
AUTHORS Kremetzki, C., Higginbotham, J., Wylie, K., Carter, J., McPherson, D.,
Warren, W., Graves, T., Mardis, E. and Wilson, R.
Gallus gallus BAC End Reads
Unpublished (2003)
Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@wustl.wustl.edu
Insert Length: 182000 Std Error: 0.00
Seq primer: Spe ATTGAGTGACACTATAG
Class: BAC ends
High quality sequence start: 422
High quality sequence stop: 487.
Location/Qualifiers
1.1434
/organism="Gallus gallus"
/mol_type="genomic DNA"
/strain="Red Jungle Fowl"
/db_xref="taxon:9031"
/clone="CH261-98P15"
/sex="female"
/cell_line="UCD001, inbred 256"
/clone_lib="CH261"
/note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
CH261 Female Chicken library - For library and clone
ordering information: <http://www.chori.org/bacpac>"

ORIGIN

Query Match 2.6%; Score 124.8; DB 28; Length 1434;
Best Local Similarity 40.6%; Pred. No. 1.8e-09;
Matches 567; Conservative 0; Mismatches 830; Indels 0; Gaps 0;

1098 ATATATGAAGACTGTGTGATCATGTGATGAACAATAGTATGATGTTCTAAATTA 1157
Db 38 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 97
1158 TATACATCAAGATTTATAGGTGTCAGATTGAGCTGCAATTAAGAACATTAATTGCA 1217
Db 98 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 157
1218 TTGCATGTGATTAATCGCTGGAATAAATTTAGTATATGCTGTGCTACTGTTATTA 1277
Db 158 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 217
1278 CTAAAGCATGATGAATTAATAACCTATATATAGCAAAATATCATTAATGATCTTC 1337
Db 218 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 277
1338 ATACATTAATGTCATCATGCTTGGAGATCTATATTTACATGTCACAGACATT 1397
Db 278 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 337
1398 CACGCATATGCTTTGGACTAGAAATAGAAAAGTAGAATATTAATACATTAATAG 1457
Db 338 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 397

Qy 1458 ATCAACAATTAAGCTTTGTAAGAACGATCTGTAACCACTGATATCTATTAGCA 1517
Db 398 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 457
1518 AAAAATTAATGTGAATCAATCAATTTGATATCTATTTCAATTTATTCATGAGATA 1577
Db 458 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 517
1578 TATCACTAGTAAGCCATATCAAACTATATCTTATGTCATCTCTTGTATTTTC 1637
Db 518 AAAAAAAAAAACCACCCACCCACCCACCCACCCACCCACCCACCCACCCACCCACCC 577
1638 AACACATGATTAATTAAGCTCAAAATATAGATTAAGTCAAGTTCACCATATATATA 1697
Db 578 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 637
1698 AGCAGTCAATACCCATGCTTAACATCACTAAGTATGTAACATATTAACATCAAACT 1757
Db 638 AAAAAAAAAAACCACCCACCCACCCACCCACCCACCCACCCACCCACCCACCCACCC 697
1758 TAAGTATATCTTAATTTCTGCTATATATTAACCAATTCATCTCTATCTCA 1817
Db 698 CAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 757
1818 TTAACCAATTTGTTATCAGATTATTAATTAAGTCAATTAACCTTTTATTTATTAAT 1877
Db 758 AAAAAAAAAAACCACCCACCCACCCACCCACCCACCCACCCACCCACCCACCCACCC 817
1878 AATATGTTGACACCTTTTCACTACATCTGATATTAAGACATTAATTTTCAATTA 1937
Db 818 AAAAAAAAAAACCACCCACCCACCCACCCACCCACCCACCCACCCACCCACCCACCC 877
1938 TAACATTAATAAAGACCAATATATATTAATTAATTAACCACTTTTACTAGTAA 1997
Db 878 AAAAAAAAAAACCACCCACCCACCCACCCACCCACCCACCCACCCACCCACCCACCC 937
1998 TATTAATATATGCTGATATTAATTTCTATCTCTGATCTGATTTCAACATTA 2057
Db 938 AAAAAAAAAAACCACCCACCCACCCACCCACCCACCCACCCACCCACCCACCCACCC 997
2058 ATATGATTAACCTGTTTTTATGACCAATTAATTTCTTAATTAATTTGCTATTAACCA 2117
Db 998 AAAAAAAAAAACCACCCACCCACCCACCCACCCACCCACCCACCCACCCACCCACCC 1057
2118 AGTTAATAATACATCTCTGCAACATATGAAGATATCATACGTCATTAATTT 2177
Db 1058 AAAAAAAAAAACCACCCACCCACCCACCCACCCACCCACCCACCCACCCACCCACCC 1117
2178 AAACCAATTTCTATATTAACCACTCTTAACATAGCAACCAATGATATCATACCT 2237
Db 1118 AAAAAAAAAAACCACCCACCCACCCACCCACCCACCCACCCACCCACCCACCCACCC 1177
2238 AAGTAAATATCCCTACAGATGATCAATTAATAAGAAAAATCGTTATGTTATTTAT 2297
Db 1178 AAAAAAAAAAACCACCCACCCACCCACCCACCCACCCACCCACCCACCCACCCACCC 1237
2298 ACATTAACAATTTTCAACCTCAACCAACCACTCAAGTTCTATTAACAATTTCA 2357
Db 1238 AAAAAAAAAAACCACCCACCCACCCACCCACCCACCCACCCACCCACCCACCCACCC 1297
2358 GAATATATATCTTCAACCAACCAATGTTATCAATTTCTATATATCATCATATCTCTTT 2417
Db 1298 AAAAAAAAAAACCACCCACCCACCCACCCACCCACCCACCCACCCACCCACCCACCC 1357
2418 TAAAGATGATCTTCAATTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2477
Db 1358 AAAAAAAAAAACCACCCACCCACCCACCCACCCACCCACCCACCCACCCACCCACCC 1417
Qy 2478 CATTAACAATTCAGA 2494
Db 1418 AAAAAAAAAAACCACCCACCCACCCACCCACCCACCCACCCACCCACCCACCCACCC 1434

FEATURES

Location/Qualifiers
 1..1373
 /organism="Pristionchus pacificus"
 /mol_type="genomic DNA"
 /strain="California"
 /db_xref="taxon:54126"
 /clone_lib="Ppa EcoRI BAC Library"
 /note="The library was generated by a partial digest of the genomic DNA with EcoRI and cloning into the BAC vector."

ORIGIN

Query Match 2.6%; Score 124.2; DB 29; Length 1373;
 Best Local Similarity 41.3%; Pred. No. 2.2e-09;
 Matches 441; Conservative 0; Mismatches 627; Indels 0; Gaps 0;

1420 AGAATATGAGAAAGGAGAAATATTAATCATTAATAGTCACAACTTAAGCTTGTGA 1479
 1356 AA 1297
 1480 AGGAAACCACTAGTAAACCACTGATATCATTAAGCAAACTTAAGTGAACCTACC 1539
 1296 AA 1237
 1540 AATTGCAATCTATTACATTTTATCATGAGATATTCCTACTAGATTAAGCCATATC 1599
 1236 AA 1177
 1600 AATCATATATCTTGTATCATCTCTGATTAATTCACACATGATTAATTAAGGT 1659
 1176 AA 1117
 1660 CAAATATAGATTCACAGCTTCACCATATATTAAGACATCAATACCATCTGA 1719
 1116 AA 1057
 1720 ACATCAACTAGTATGATTAATTAATTAATTAATTAATTAATTAATTAATTCCT 1779
 1056 AA 997
 1780 AGCTAATATATTAACAAATTCCTATCTCAATTAACCAATTTGATGAGATTA 1839
 936 AA 937
 1840 TAAATATGCAATTAACCTTTTACTTATTAATTAATTAATTAATTAATTAATTC 1899
 936 AA 877
 1900 CTACATCATGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1959
 876 AA 817
 1960 CATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2019
 816 AA 757
 2020 AATAATTCCTATCATCTTGAATTTCAACATTAATTAATTAATTAATTAATTTTAT 2079
 756 AA 697
 2080 GCCCTTGTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2139
 656 AA 637
 2140 AACATATGAGAAAGATACATACACGCAATTAATTAATTAATTAATTAATTAATTA 2199
 636 AA 577
 2200 CTTACTTAACATAGGAAACCAATGATATCATCTTAAGTAAAGATATCCCTACAGAT 2259
 576 AA 517
 2260 GTATCAATATTAAGAAATATGATATTAATTAATTAATTAATTAATTAATTAATTC 2319

Db 516 AA 457
 Qy 2320 CAACAACAAACTCAAGTTCTATTATTAACAAATTCAGAGATATTAATTAATTAAT 2379
 Db 456 AA 397
 Qy 2380 ACATAGTATCAATTTCTATTAATCAATCATATCTTTTAAAGTGAATTCATTC 2439
 Db 396 AA 337
 Qy 2440 ATACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2487
 Db 336 AA 289

RESULT 21
 BQ455897/C 546 bp mRNA linear EST 29-MAY-2002
 LOCUS k22401.y1 Dicrofilaria immitis adult PAMP: v1 Dicrofilaria immitis
 DEFINITION cDNA 5' similar to TR:P90551 P90551 GLYCEROL-3-PHOSPHATE
 DEHYDROGENASE, mRNA sequence.
 ACCESSION BQ455897
 VERSION BQ455897.1 GI:21259002
 KEYWORDS EST.
 SOURCE Dicrofilaria immitis
 ORGANISM Dicrofilaria immitis
 Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioides;
 Onchocercidae; Dicrofilaria.

REFERENCE
 1 (bases 1 to 546)
 McCarter, J., Clifton, S., Chapell, B., Page, D., Martin, J.,
 Wylie, T., Dante, M., Marra, M., Hillier, L., Kucaba, T., Theising, B.,
 Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Franklin, C.,
 Tsagaris, V., Ronko, I., Kennedy, S., Maguire, L., Beck, C.,
 Underwood, K., Steptoe, M., Allen, M., Person, B., Swallow, T.,
 Harvey, N., Schurk, R., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
 McCann, R., Waterston, R. and Wilson, R.
 The Washington Univ. Nematode EST Project, 1999
 The Washington Univ. Nematode EST Project, 1999
 Contact: McCarter JP
 The Washington Univ. Nematode EST Project, 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.edu

COMMENT
 The library was constructed by Claire Murphy and Dr. James McCarter
 at Washington University, St. Louis. Adult nematodes were harvested
 from infected dogs by Dr. Prema Arasu of North Carolina State
 University, Raleigh, NC (Prema.Arasu@ncsu.edu). Total RNA was
 isolated by Merry McAlaird of Divergence, Inc., St. Louis, MO. DNA
 Sequencing by: Washington University Genome Sequencing Center

FEATURES
 source
 Possible reversed clone: similarity on wrong strand
 Seq primer: -40RP from Gibco
 High quality sequence strop: 418.
 Location/Qualifiers
 1..546

/organism="Dicrofilaria immitis"
 /mol_type="mRNA"
 /db_xref="taxon:6287"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="Dicrofilaria immitis adult PAMP1 v1"
 /note="Vector: PAMP1 (Gibco), Site 1: NotI, Site 2: SalI;
 The library was constructed by Claire Murphy and Dr. James
 McCarter at Washington University, St. Louis. The cDNA was
 made by using Dynabead Oligo-dT priming (Dyna1). PCR based
 library using a modified protocol from the SMART PCR cDNA
 Synthesis Kit from Clontech. Directionally cloned into the
 UDG sites of PAMP1. Adult nematodes were harvested from
 infected dogs by Dr. Prema Arasu of North Carolina State
 University, Raleigh, NC (Prema.Arasu@ncsu.edu). Total RNA

ORIGIN

was isolated by Merry McIaird of Divergence, Inc., St. Louis, MO."

Query Match 2.6%; Score 123.4; DB 13; Length 546;
Best Local Similarity 58.6%; Pred. No. 4.1e-09;
Matches 214; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

QY 953 AAAGATGCAATTAATCACTCAATTAATCCAGTAATAGCAGAAATTTTACAA 1012
DB 546 AAAGATGCAATTAATCACTCAATTAATAGCAAGAAATTAAGATTTTACCT 487
QY 1013 TATATCCAAATTTTACTCTCTGTCCTCAAGTTTGTCTAAAGAAATTTGCAACATCTT 1072
DB 486 AACACCCCTGTCCTATTTTTCAGGTCCTAGCTTTCGATGGAAGTCGCAAAAAATCG 427
QY 1073 CCTTGTATATAGTACTCTGCTGATATATAAGAACTTGATCAATGATGAACA 1132
DB 426 CCTTATCAATGCTTCTGATGATCAAACTATACCTAGGCTCAAGTTGATGACAA 367
QY 1133 ATATGATATGATGTTCTAAATATATATCAATCAAGATTTATAGGTACAGATTGCA 1192
DB 366 CTATACCAAGAAATTAATCAAGTTTATAGTATGATTTATAGGGGTGCAATTTTGT 307
QY 1193 GCTGATTAAGACATTAATTCATTCATGATGTAATGCTGGAATTAATTTAGT 1252
DB 306 GCGACATTAAGACATTTTCTGATGATGATGATGATGATGATGATGATGATGATG 247
QY 1253 AATATGCTGTTGCTACTGTTATTAATTAAGGATGAATGAATTAACAATATATA 1312
DB 246 TACAAATCCACGACGATGTTAGGCGATTAATTAATGACATGACATCTTAATTC 187
QY 1313 GCATA 1317
DB 186 AGATA 182

RESULT 22
CG758143 1297 bp DNA linear GSS 24-OCT-2003
LOCUS CG758143/c
DEFINITION P053-3-B11.zb Ppa EcORI BAC Library Pristionchus pacificus genomic.
ACCESSION CG758143
VERSION CG758143.1 GI:37987385
KEYWORDS GSS.
SOURCE Pristionchus pacificus
ORGANISM Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
Nemodiplogasteridae; Pristionchus.
1 (bases 1 to 1297)
Strihvasan,J., Shinz,W., Jesse,T., Wiggers-Perebolle,L., Jansen,K.,
Bunjer,J., van der Meulen,M. and Sommer,R.J.
An integrated physical and genetic map of the nematode Pristionchus
pacificus
Mol. Genet. Genomics 269 (5), 715-722 (2003)
JOURNAL MEDLINE
PUBMED 22835951
COMMENT 12884007
Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
Class: BAC ends.

FEATURES

source
1..1297
Location/Qualifiers
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_id="Ppa EcORI BAC library"
/note="The library was generated by a partial digest of

ORIGIN

the genomic DNA with EcoRI and cloning into the BAC vector."

Query Match 2.6%; Score 123; DB 29; Length 1297;
Best Local Similarity 40.9%; Pred. No. 3.5e-09;
Matches 441; Conservative 0; Mismatches 636; Indels 0; Gaps 0;

QY 1276 AACATAGCATGATGAATTAATTAACATATATATAGCAAAATCATTCATGATCT 1335
DB 1290 AA 1231
QY 1336 TCATACATTAATGTCATCATGCTTGGAGATCTTAATATTAACATGACACAGACA 1395
DB 1230 AA 1171
QY 1396 TTCACGCAATATGCTTTGGACTAGAAATAGCAAAAGTGAATATATATCATTAAT 1455
DB 1170 AA 1111
QY 1456 AGATCAACCTTAAGCTTGTGAGAAACCACTACTGTAAACCACTGATATCATTAAC 1515
DB 1110 AA 1051
QY 1516 AAAAAAAAACTTAATGTAACATACCAATTTGATATTTTACATTTTACATGAGA 1575
DB 1050 AA 991
QY 1576 TATATCACTAGTAAAGCATATCAACATATATATCTTATGCTATGCTTTGATAT 1635
DB 990 AA 931
QY 1636 TCAACCAATGATTAATTAACGCTCAAAATATGATCAACCTTACCACTATATATA 1695
DB 930 AA 871
QY 1696 AAGCATCAATATACCATGTTTATACATCAATAGTATGATATTTTACATACAA 1755
DB 870 AA 811
QY 1756 CTATAGTATATCTTAATTTCTTACTATATATATTAACAAATTCCTTCTTCTAT 1815
DB 810 AA 751
QY 1816 AATTAACAATGTTTATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 1875
DB 750 AA 691
QY 1876 ATATATGTTGCACTTTTCACTACATCATGATATATGAACATATTTTACATTA 1935
DB 690 AA 631
QY 1936 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1995
DB 630 AA 571
QY 1996 AATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2055
DB 570 AA 511
QY 2056 AATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2115
DB 510 AA 451
QY 2116 AAAGGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2175
DB 450 AA 391
QY 2176 TTAACCAATTTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2235
DB 390 AA 331
QY 2236 CTATAGTATATCTTCAATGATGATCAATTAATTAAGAAATTAATTAATTTT 2295

Dd	330	AAA	271
Oy	2296	ATGACAAATTACAAAATTTCCACTTGACAACAGCAAACTCAAACTTCTATTATACAA	2352
Dd	270	AAA	214
RESULT 23			
LOCUS	CG747324/C		
DEFINITION	CG747324	1211 bp	DNA linear GSS 24-OCT-2003
KEYWORDS	P040-4-D10.za Ppa EcORI BAC Library Pristionchus pacificus genomic.		
ACCESSION	CG747324		
VERSION	CG747324.1	GI:37968250	
SOURCE	GSS:		
ORGANISM	Pristionchus pacificus		
REFERENCE	Pristionchus pacificus		
AUTHORS	Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;		
TITLE	Neodiplogasteridae; Pristionchus.		
JOURNAL	1 (bases 1 to 1211)		
MEDLINE	Srinivasan,J., Siriz,W., Jesse,T., Wiggers-Perebole,L., Jansen,K.,		
COMMENT	Bunjer,U., van der Meulen,M. and Sommer,R.U.		
	An integrated physical and genetic map of the nematode Pristionchus		
	pacificus		
	Mol. Genet. Genomics 269 (5), 715-722 (2003)		
	Contact: Sommer RJ		
	Evolutionary Biology		
	Max-Planck-Institute for Developmental Biology		
	Spemannstr. 37-39, Tuebingen D-72076, Germany		
	Tel.: 00497071601371		
	Fax: 00497071601498		
	Email: ralf.sommer@uebingen.mpg.de		
	Classes: BAC ends.		
FEATURES			
source	location/Qualifiers		
	1..1211		
	/organism="Pristionchus pacificus"		
	/mol_type="genomic DNA"		
	/strain="California"		
	/db_xref="taxon:54126"		
	/clone_lib="Ppa Scori BAC library"		
	/note="The library was generated by a partial digest of		
	the genomic DNA with EcoRI and cloning into the BAC		
	vector."		
ORIGIN			
Query Match	2.6%; Score 122.8; DB 29; Length 1211;		
Best Local Similarity	39.9%; Pred.No. 3.8e+09;		
Matches 427; Conservative	0; Mismatches 643; Indels 0; Gaps 0;		
Oy	1496	AAACCACTGATATCATTTGCAGCAAAAAAATTATGTGAACATGCCAATTTGCATATCTATT	1555S
Dd	1209	AAA	1150C
Oy	1556	TACAATTATTACAGAGAATATATCTAGATTAAGCCATATCAACATATTACTTAG	1615S
Dd	1149	AAA	1090D
Oy	1516	TCTATTCATCCTTTGATAATTTCAACACCAATGATAATTAACGGTCAAAAAATRTGAATAC	1675S
Dd	1089	NAAA	1030D
Oy	1676	AACGTTACACCATATATTAAGAACAGTCAAATACCATCGTTACATCCAATAGATAT	1735S
Dd	1029	AAA	970D
Oy	1736	GTACATATTTCAATACAACTTAAGATTAATCTTAATTTCCAGCAATATTAATAAC	1795S
Dd	969	AAA	910D
Oy	1796	AAATTCCTATCTCATCTCAATTAACCAATTGTTATCAGATTATTAATAGCAATTA	1855S

[illegible]

Email: submissions@watsen.wustl.edu
 Insert Length: 182000 Std Error: 0.00
 Seq primer: Sp6 ATTAGGTGACACTATAG
 Class: BAC ends
 High quality sequence start: 34
 High quality sequence stop: 261.

FEATURES

source

1.1272
 Location/Qualifiers
 /organism="Gallus gallus"
 /mol_type="genomic DNA"
 /strain="Red Jungle Fowl"
 /db_xref="taxon:9031"
 /clone="CH261-19L17"
 /sex="female"
 /cell_line="UCD001, inbred 256"
 /clone_lib="CH261"
 /note="Vector: pRABAC2.1; Site 1: EcoRI, Site 2: EcoRI;
 CH261 Female Chicken library - for library and clone
 ordering information: http://www.choxi.org/bacpac"

ORIGIN

Query Match 2.6%; Score 121.8; DB 28; Length 1272;
 Best Local Similarity 47.7%; Pred. No. 5.3e-09;
 Matches 407; Conservative 0; Mismatches 444; Indels 3; Gaps 2;

1604 ATATATCTAGTCTATCATCTCTGATATTTCAACA-CAATGATTAATAAGGTCAA 1662
 1236 ATATATATATAGTTAATATATATATATATATATATATATATATATATATATAT 1177
 1663 AAATAGATGCAACGTCACCATATATATATATATATATATATATATATATATATAT 1722
 1176 TAAAT 1117
 1723 TCCAACTAGTATGAT 1782
 1116 AAAAAAT 1057
 1783 TAT 1842
 1056 AT 997
 1843 AT 1902
 996 AT 937
 1903 CATCATGAT 1962
 936 TAT 879
 1963 AT 2022
 878 AT 819
 2023 AAATCTATCATCTGAT 2082
 818 TAT 759
 2083 CATCTGCTTAT 2142
 758 TAT 699
 2143 AT 2202
 698 AT 639
 2203 ACTTAAAT 2262
 638 TTAAT 579
 2263 TCAAT 2322
 578 AAAAAAT 519

QY 2323 CAACAACTCAAGTCTAT 2382
 DB 518 TAAAT 459
 QY 2383 TACTATCAATCTCTAT 2442
 DB 458 ATTTTAT 399
 QY 2443 CATTAAT 2456
 DB 398 TAT 385

RESULT 25

CG753854/c

LOCUS

DEFINITION

P049-1-D06.2a Ppa EcoRI BAC library Pristionchus pacificus genomic

ACCESSION

CG753854

VERSION

CG753854.1 GI:37978755

KEYWORDS

GSS

SOURCE

PRISTIONCHUS PACIFICUS

ORGANISM

PRISTIONCHUS PACIFICUS

REFERENCE

1 (bases 1 to 1528)

Srinivasan, J., Sinc, W., Jesse, T., Wiggere-Perreolte, L., Jansen, K.,

Buntjer, J., van der Meulen, M., and Sommer, R.J.

An integrated physical and genetic map of the nematode Pristionchus

pacificus

Mol. Genet. Genomics 269 (5), 715-722 (2003)

CONTACT: Sommer, R.J.

12884007

COMMENT

Evolutionary Biology

Max-Planck-Institute for Developmental Biology

Spermannstr. 37-39, Tuebingen D-72076, Germany

Tel: 00497071601371

Fax: 00497071601498

Email: ralf.sommer@tuebingen.mpg.de

Class: BAC ends.

Location/Qualifiers

1.1528

/organism="Pristionchus pacificus"

/mol_type="genomic DNA"

/strain="California"

/db_xref="taxon:54126"

/clone_lib="Ppa EcoRI BAC library"

/note="The library was generated by a partial digest of

the genomic DNA with EcoRI and cloning into the BAC

vector."

vector."

vector."

vector."

vector."

vector."

vector."

vector."

vector."

vector."

vector."

vector."

vector."

vector."

ORIGIN

Query Match 2.6%; Score 121.8; DB 29; Length 1528;
 Best Local Similarity 37.7%; Pred. No. 5e-09;
 Matches 441; Conservative 0; Mismatches 729; Indels 0; Gaps 0;

QY 1276 AACTAAGGCAATGAT 1335
 DB 1485 AANNNNNANNNNNANNNNNANNNNNANNNNNANNNNNANNNNNANNNNNANNNNA 1426
 QY 1336 TCATACATTAATTTGTCATCATGCTTGAGATCTAATATTATATATATATATATATAT 1395
 DB 1425 AANNNNNANNNNNANNNNNANNNNNANNNNNANNNNNANNNNNANNNNNANNNNA 1366
 QY 1396 TTCACCAATATAGCTTTTGCACTAGAAATAGGAAATGAAATATATATATATATATAT 1455
 DB 1365 AANNNNNANNNNNANNNNNANNNNNANNNNNANNNNNANNNNNANNNNNANNNNA 1306
 QY 1456 AGATCAACAACCTTAAGCTTTGTAAGGAAACAGTATCTGTAATATATATATATATAT 1515
 DB 1305 AAAAAANNNNNANNNNNANNNNNANNNNNANNNNNANNNNNANNNNNANNNNAAN 1246

Oy	1516	AAAAAACTTATGAGACACTCCATTGGCACTACTTTTACAATTTATTAACATGAGA	1575
Dd	1245	NANNAAAAAAAAAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNA	1186
Oy	1576	TATATCAGTATAAGCCATATCAACAATATTATCTTAGTCATCATCTTGTAATT	1638
Dd	1185	AAAAAAAAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNA	1126
Oy	1636	TCACACATGTATAATTAAGCGTCAAAATTTAGATATCAACGTTCCCATATATA	1695
Dd	1125	AAAAAAAAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNA	1066
Oy	1636	AAAGAGTCAAATACCATCGTTAACATCCACTAGATGTAAACATTTACATACAA	1755
Dd	1065	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANNAANNAANNAANNA	1006
Oy	1756	CTTAGTATTAAGTTAAATTCCTAGCTATATATTAACAATTCCTATCTCATCTC	1815
Dd	1005	ANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNA	946
Oy	1816	AATTACCAATTTGTTATGATTTATAATTAAGTCAATTAACCTTTCTTATTATA	1875
Dd	945	AAAAAAAAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNA	886
Oy	1876	ATAATGTGTGCACATTTTTCTACTACATCATGTATATAATGAACATATTTACATA	1935
Dd	885	AAAAAAAAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNA	826
Oy	1936	AATACTATATAAAGAACCAATCATATTTAAATTACTAAAACCTATTTCTACGT	1995
Dd	825	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANNAANNAANNAANNA	766
Oy	1996	AATATTACATGAGTCTCTGACATATATAATTTCTATCATCTTGACATATTTCAA	2055
Dd	765	ANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNA	706
Oy	2056	AAATGTGTAACCTGTTTTTTTAAAGCCATTAGTTCTTAATATAATGCTAATTACCA	2115
Dd	705	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANNAANNAANNAANNA	646
Oy	2116	AAAGGTAAATACACATTCCTGCAACATATGAAGAATACATACAGTCAATATA	2175
Dd	645	NNNAAANNAANNA	586
Oy	2176	TTAAACAAATTTCTPATTAACCCCTACTCTAAACAATAGGAACCAATAGATATCATC	2235
Dd	585	AAAAAAAAAAAAAAAAANNAANNAANNAANNAANNAANNAANNAANNA	526
Oy	2236	CTAAGTAAAGATATCCTTACAGCATGTATCAATATATAAGAAAATCGTTATGTATTT	2295
Dd	525	AAAAAAAAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNA	466
Oy	2296	ATACATTTACAAATTTCCATCTTCAACAACAAATCTAAAGTTCTATTAACAAATTC	2355
Dd	465	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANNAANNAANNAANNA	406
Oy	2356	AAGAATTATATATCTTCAACACACATAGTTATCAATTTCTATATACATCATATCCTT	2415
Dd	405	AAAAAAAAAAAAAAAAAAAAAAAAANNAANNAANNAANNAANNAANNNNNNN	346
Oy	2416	TTTAAAAAGATGACTTCACATTCATACAT	2445
Dd	345	ANNAANNAANNNNNNNNNAACCATTCAT	316

RESULT	26
CG744812/c	
LOCUS	CG744812
DEFINITION	P037--B03.ya Ppa ECORI 1353 bp DNA linear GSS 24-OCT-2003
ACCESSION	CG744812
VERSION	CG744812.1
KEYWORDS	GI:37965680

ORIGIN	SOURCE	FEATURES
ORGANISM	Pristionchus pacificus	
REFERENCE	Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida; Neodiplogasteridae; Pristionchus.	
AUTHORS	1 (bases 1 to 1353)	
TITLE	Strinivasan,V., Sinz,W., Jesse,T., Wiggers-Derebolte,U., Jansen,K., Buntjer,J., Van der Neulen,M. and Sommer,R.J.	
JOURNAL	An integrated physical and genetic map of the nematode Pristionchus pacificus	
MEDLINE	Mol. Genet. Genomics 269 (5), 715-722 (2003)	
PUBMED	22835951	
COMMENT	12884007	
	Contact: Sommer RJ	
	Evolutionary Biology	
	Max-Planck-Institute for Developmental Biology	
	Spemannstr. 37-39, Tuebingen D-72076, Germany	
	Tel.: 00497071601371	
	Fax: 00497071601498	
	Email: ralf.sommer@uebingen.mpg.de	
	Class: BAC ends.	
	Location/Qualifiers	
	1..1353	
	/organism="Pristionchus pacificus"	
	/mol_type="genomic DNA"	
	/strain="California"	
	/db_xref="taxon:54126"	
	/clone_lib="Ppa Scott BAC Library"	
	/note="The library was generated by a partial digest of the genomic DNA with EcoRI and cloning into the BAC vector."	
Query Match	2.5%; Score 120.6; DB 29; Length 1353;	
Best Local Similarity	39.6%; Pred. No. 7.9e-09;	
Matches	422; Conservative 0; Mismatches 643; Indels 1; Gaps 1;	
QY	1420 AGAATAGGAAAGGTGAAATATATAATACATTATATAGATCACAACCTAAAGCTTGTGA	1479
DB	1270 AA	1211
QY	1480 AGGAACGACGTCTGTAAAAACGACGATATCATTTGCAAAAAAATAATATGTGAACTAC	1539
DB	1210 AA	1151
QY	1540 AATTGCATATCTATTTACATTATTTACATGCAATATATCTACTAGTAAAGCATATC	1599
DB	1150 AA	1091
QY	1600 AAACATATCTTAGTCTATCATCTCTTGATATTT-CAACCAATGATATAATAAACGG	1658
DB	1090 AA	1031
QY	1659 TCATAAATATATGATATCAACGTTTACACACATATATAAAAGAGTCAAAATCCCATGTT	1718
DB	1030 AA	971
QY	1719 AACATCCAACTAGTATGTAACATATTTACATTAACAACCTTAACTATTAATTTCC	1778
DB	970 AA	911
QY	1779 TAGCTAATTTATATAACAATTTCTATCTCTCTCATTTAACCAATTTGTTACAGATT	1838
DB	910 AA	851
QY	1839 ATAAATAGTCAATTAACCTTTACTATTATTAATATATATATGTTTGACACTTTTC	1898
DB	850 AA	791
QY	1899 ACTACATCATGTATATATAGACATATTTTACATTAATAATACATTAATAAAGACCAAT	1958
DB	790 AA	731
QY	1959 ACATATTTAATTTACTAAACACCTATTACTACGTATACGTATATACATAGATGTCCTGACA	2018

Df		720	AAA	671
Oy		2019	TGATTAATTCTACATCCTTGCAATTTCAAACATAAAATGTGAATAGTCTTTTTTTTA	2078
Df		670	AA	611
Oy		2079	TGCCCATYAGTCCTTAATTTAAATGGTCTAATTACCAGAAAAGTTAAATACACTTCCTG	2138
Df		610	AA	551
Oy		2139	CACAAATATGAGAATATCAATCACCGTCGAATTAATTAACAATTTCTATATTAAC	2198
Df		550	AAMNAAA	491
Oy		2199	ACCTACTCTAAAATAGAACCAATAGTATATCATACCTAAGTAAGATATCCCTACAGA	2258
Df		490	ANNAANAANNAAA	431
Oy		2259	TGTATCAAAATATAGAAAAATCGTATGTATTTATTTACATTAACAAATTTCCAAT	2318
Df		430	AA	371
Oy		2319	TCACACACAAAATCTCAAGTCTATATTAACAAATTCAGATTTATATATCTTACACCA	2378
Df		370	AA	311
Oy		2379	CACATAGTTATCAATTTCTATATACATCATATCCTTTTAAAAAGATGACTTCACATT	2438
Df		310	AA	251
Oy		2439	CATACATAATATATATATACATGCATGTTTATTCAGAAACATTAACA	2484
Df		250	NNANNA	205
RESULT 27				
LOCUS	CC253231	1277 bp	NDA	linear
DEFINITION	CH261-180N11.RM.1 CH261 Gallus gallus genomic clone CH261-180N11,			
ACCESSION	CC253231			
VERSION	CC253231.1	GI:30589981		
KEYWORDS	GSS.			
SOURCE	Gallus gallus (chicken)			
ORGANISM	Gallus gallus			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Archosauria; Aves; Neognathae; Galliformes; Phasianidae;			
	Phasianinae; Gallus.			
	1 (bases 1 to 1277)			
REFERENCE	Karenikhi, C., Higsibotham, J., Wylie, K., Carter, J., McPherson, J.,			
AUTHORS	Warren, W., Graves, T., Mardis, E. and Wilson, R.			
JOURNAL	Gallus gallus BAC End Reads			
COMMENT	Unpublished (2003)			
	Contact: Richard K. Wilson			
	Genome Sequencing Center			
	Washington University School of Medicine			
	Email: submissions@watson.wustl.edu			
	Insert Length: 182000 Std Error: 0.00			
	Seq primer: RM1 TAGCAGTCACATATGGAGA			
	Class: BAC ends			
	High quality sequence start: 24			
	High quality sequence stop: 85.			
FEATURES	Location/Qualifiers			
source	1..1277			
	/organism="Gallus gallus"			
	/mol_type="genomic DNA"			
	/strain="Red Jungle Fowl"			
	/db_xref="taxon:9031"			
	/clone="CH261-180N11"			
	/sex="female"			
	/cell_line="UCD001, inbred 256"			
	/clone_id="CH261"			

ORIGIN

/note="Vector: pTARBAC2.1; Site 1: EcoRI, Site 2: EcoRI
CH261 Female Chicken library - For library and clone
ordering information: <http://www.chori.org/bacpac>"

Query Match	2.5%	Score 120.4;	DB 28;	Length 1277;
Best Local Similarity	46.8%;	Pred. No. 8.6e-09;		
Matches 491;	Conservative 0;	Mismatches 546;	Indels 13;	Gaps 3;

Qy	1438	AAAAATAAATAACATTAAATAGTCTCAACACCTAAAGCTGGTAGAGAAACAGACTCGTAA	1497
Db	231	AATATATATTAATTAATTATTATTAATAATATAAATATATAACATATATAATATA	290
Qy	1498	ACCATGTATCATTAGCAAAAAACCTTAATGTGAACTACCAATTTGCAATCTATTTA	1557
Db	291	ATTAAATTAATAATATACATATTTATATATTTCAAAAAATTAAAAATTAATATATAATAT	350
Qy	1558	CAATTTATCATGAGAAATATATCACTAGATTAAGGCAATATCAACATATATCTGTGTC	1617
Db	351	AATATATATTAATATATTAATTAATTTATATATATATAAATAATATATATTAATA	410
Qy	1618	TATCATCTTGATTAATTTCAACAAATGATTAATAAACGGTCAAAAATATGAAATACAA	1677
Db	411	TAAAAATATATATTAATAATATTTTATTAATTAATAAAAAATTAAATATATTTTATAT	470
Qy	1678	CGTCCACACATATATAAAGGAGCAAAATCCATCGTTAATACATCACTAGATGT	1737
Db	471	TTTTATATTAATATATTAATTAATTTTATATATATATAAATAATATAATTAATACT	530
Qy	1738	AAC-----ATATTCATTAACAACCTTAAGTATACCTTAATAATTCCTAGCTAATATA	1798
Db	531	ATATATTTTAAATATTTAAATATATATTTTAAATTAATAAATAATTAATATATAATATATT	590
Qy	1791	TAAACAATTCCTATCTCCATCCATTTAACCATGTGTATCGATTTATATAAAGCA	1856
Db	591	AAATAAAAATTAATTTAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTA	650
Qy	1851	AATTAACCTTTACTTATATATAAATAATATGTTGACACTTTTCACTACATCATGT	1910
Db	651	AATTAATATTTAAATATATATATAAATAATTAATTAATTAATTAATTAATTAATTTT	710
Qy	1911	A---TATATAGACATATTTTTCATATAAATACTAATTAATAAGAACCAATACATATTT	1967
Db	711	AATTTTAAATAATTAATTAATTAATTTTAAAAATTTTATTAATTAATTAATTAATTA	770
Qy	1968	AAATTAATAAACCCATTTTACTAGCTAATATATACATAGTGCCTGACATATAATTT	2027
Db	771	ATTTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	827
Qy	2028	CTATCATCTTGACATATTTCAACAACATTAATATGATTAACGTGTTTTTATGCCCCATTA	2087
Db	828	ATAAAAATTTATAPATATTTTATATTTATTAATAAAAAAATTAATAATATATATTAATA	887
Qy	2088	GTCTTATATAAATTTGCTAATTTCCCAAAAAGTTAAATAACACATCTCGCAACAATAT	2147
Db	888	TTTAACATTAATTAATTAATTTTAAATTAATTAATTTTAAATTAATAAATTAATTAATTTT	947
Qy	2148	GAGAGAGATACAATACACGCTCAATAAATTTAAACAAATTTCTATATATACCACTACTCT	2207
Db	948	ATATTAATTAATTAATTAATTTTATATTAATTAATTAATTAATTAATTAATTAATTTATTA	1007
Qy	2208	AAACATAGGAACAATGTATATCATACCTAAGTAGATATGCTCCCTAGAGATGATCAAA	2267
Db	1008	TAAAAAATAATTAATTAATTTTATATATTTTCTATATTTAATTAATTAATTAATTAATTA	1067
Qy	2268	TAAATAAGAAAAATCGTTATGATTAATTTATACAAATTCAAAAATTTCCAACTTCAACACA	2327
Db	1068	AAATTTAATTTATTAATTAATTAATTAATTAATTAATTTAAATAATTAATTAATTAATTA	1127
Qy	2328	AAACTCAAAAGTCTATTAATACAAATTCGAAGATTTATATATCTTCCACCAACATGTGT	2387
Db	1128	TBAATTAATTTAAATAATTTATATATATTTTATTAATAAATAATATATTAATTAATTAATTA	1187

[illegible]

FEATURES

SOURCE

	Query Match	2.5%	Score 120.2	DB 29	Length 1433
	Best Local Similarity 41.4%		Pred. No. 8	9e-09	
	Matches 457	Conservative 0	Mismatches 646	Indels 1	Gaps 1
Qy	1275	TAACATAAGCATGATGAAATTTAAACATATATATAGCAAAAAATCATTCATATAGATC	1334		
Db	1423	TAAATTAATAAAAAATTAANAATAATTAATATTTATTAATAANAATAATTAATTAATAANAAN	1364		
Qy	1335	TTCATACATTTATGTGTCATCATGCTCTTGAGATCTAATTTATACATGTGTCAACAGAC	1394		
Db	1363	TAAATAAANNNTAANAANAABAAAAAABAAAAAABAAAAAABAAAAAABAAAAAABAAAAA	1304		
Qy	1395	ATTACAGCAATATGGCTTTTGGACTAGAAATAGAAAAGGTGAATATTAATTCATTAA	1454		
Db	1303	AAAAAABAAAAAABAAAAATTAANAANAABAAAAAABAAAAAABAAAAAABAAAAAABAAAAA	1244		
Qy	1455	TAGATCAAACTCTAAAGCTTGTGAAGAAACAAGTACTGTAAAAACAATGATCATTAG	1514		
Db	1243	AAANAANAABAAAAAABAAAAAANNANNNANNAANNTAAAAAABAAAAAABAAAAAABAAAAA	1184		
Qy	1515	CAAAAAAATTAATGTGAACCTACCAATTG-CATATCTATTTTCAATTTTATTCATGAG	1573		
Db	1183	AAAAAABATTAANAANAABAAAAAABAAAAAABAAAAAABAAAAAABAAAAAABAAAAA	1124		

QY	1574	AAATATACAGATTAAGCAATATCAAACTATTGCTAGTCATCATCTTGATPA	1633
Db	1123	AA	1064
QY	1634	TTTCACACAAATGATTAATAAAGCGTCAAAAATTTAGATTCGAACGTTACCA	1639
Db	1063	AA	1004
QY	1694	AAAAAGCGTAATACCACATCGTTAACTCACTAAGTATGPAACATATTACATA	1753
Db	1003	ANNNAA	944
QY	1754	AACCTAAGTATACCTTAAATTTCTTGCTATATTATATTAACAAATTCCTATC	1813
Db	943	AA	884
QY	1814	TCAATTAACCAATGTTATCAGATTATTAATTAAGTCAATTAACCTTTACTTA	1873
Db	883	NNNNAA	824
QY	1874	AAATATATGTTGTCACACTTTTCTCACTCATCTGATATATGAACATATTTTAC	1933
Db	823	AA	764
QY	1934	AAATTAATCTAATAAAGAACCAATCATATATTAAATTACTAAACCCATTTACTAC	1993
Db	763	AA	704
QY	1994	GTAATATTCTAGATGTCGACATATAATTCTATCTTCCTTGACATTTCCACACA	2053
Db	703	AA	644
QY	2054	TAAATATGATAAAGCTGTTTTTTATGCGCCATTAGTCTTATAAATTGCTAATTA	2113
Db	643	AA	584
QY	2114	AAAAAGGTTAAATATACCATCTTCGACACATATGAAGAGATCAATACGCTCAATA	2173
Db	583	AA	524
QY	2174	AATTAACAATTTCTATATTAACCACTTACTTAACATAGAACCAATAGTATATCAT	2233
Db	523	AA	464
QY	2234	ACCTAAGTATGATTCCTCTACAGATGATATCAATTAATAAGAAAAATCGTAGTAT	2293
Db	463	AA	404
QY	2294	TTATACATTAACAATAATTTCCACTTCACACACAAAACTCAAGTTCTATTATACAAT	2353
Db	403	AA	344
QY	2354	TCAAGATTATATTATTACTTACCA 2377	
Db	343	AAAAAAAAAAAAAAAAAATCCAATACTA 320	
RESULT 29			
CG750135	1592 bp	DNA	linear
LOCUS	P044-3-D09.2a	Ppa Ecobi BAC library	Pristionchus pacificus genomic,
DEFINITION		genomic survey sequence.	
ACCESSION	CG750135		
VERSION	CG750135.1	GI:37971278	
KEYWORDS	GSS.		
SOURCE	Pristionchus pacificus		
ORGANISM	Pristionchus pacificus		
REFERENCE			
AUTHORS	Stintzi, J., Sins, W., Jesse, T., Wiggers, P., Perle, L., Jansen, K.,		
	Buntjer, J., van der Meulen, W., and Sommer, R. J.		
	An integrated physical and genetic map of the nematode Pristionchus		
TITLE			

JOURNAL
MOL. Genet. Genomics 269 (5), 715-722 (2003)
MEDLINE
22835951
PUBMED
12884007
COMMENT
Contact: Sommer RU
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
Class: BAC ends.

FEATURES

source
Location/Qualifiers
1..1592
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="EcoRI BAC Library"
/note="The library was generated by a partial digest of
the genomic DNA with EcoRI and cloning into the BAC
vector."

ORIGIN

Query Match 2.5%; Score 119.6; DB 29; Length 1592;
Best Local Similarity 42.2%; Pred. No. 1.1e-08;
Matches 497; Conservative 0; Mismatches 680; Indels 2; Gaps 1;
QY 1420 AGAATAGAGAAAGAGAGAAATTAATTCATTAAATAGATCAACCTTAAGCTTGTGA 1479
DB 226 AAAAAAAAAAANN 285
QY 1480 AGGACCACTACTGTAACCACTGATATCATTAGCAAAAACTTAATGTAAGTACC 1539
DB 286 AAAAAAAAAAANN 345
QY 1540 AATTGATATCTATTTATTAATTTATCATGAGATATATCATGATTAAGCATATC 1599
DB 346 AAAAAAAAAAANN 405
QY 1600 AATCATATATCTTATGCTATCATCTTGTATTAATTCACCAATGATTAATAACGGT 1659
DB 406 AAAAAAAAAAANN 465
QY 1660 CAAAAATATAGATCAAGCTTCACCACTATATTAAGCAAGTAAATACCCATCGTA 1719
DB 466 AAAAAAAAAAANN 525
QY 1720 ACATCCAACTAAGTATGTAACATATTAACAACCTTAAGTATTAATTAATTTCT 1779
DB 526 AAAAAAAAAAANN 585
QY 1780 AGCTAATTAATTAACAATTCCTATCTCTCAATTAACCAATGTTATCAATTA 1839
DB 586 TTTTATTTTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTA 645
QY 1840 TAAATAGCAATTAACCTTTTACTTATTTAATAATATGTTGTCACACTTTTCA 1899
DB 646 TTAATATTAATTAATTAATTAATTAATTAATTAATTTTAAATTTTAAATTA 705
QY 1900 CTACATCATGTATTAAGCAATATTTTACATATAAATACTAATTAAGAACATA 1959
DB 706 TTAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATA 765
QY 1960 CATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2019
DB 766 TTAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTT 825
QY 2020 AATTAATTCATCATCTTGTACATATTCACAAATTAATTAATTAATTAATTTAT 2079
DB 826 TTAATATTAATTAATTAATTTTATTAATTAATTTTATTAATTAATTAATTAATTT 885
QY 2080 GCCATAGTCTTATTAATTAATTTGCTAATTAACCCAAAAAGTTAAATTAACATTCCTGC 2139

DB 886 TTTATTTATTAATTTATTAATTTATTAATTTATTAATTTATTAATTTATTAATTTA 945
QY 2140 AACATATGAGAGATACATACATACATCAATTAATTAATTAATTAATTTCAATTAACCA 2199
DB 946 ATTAATTTATTAATTTATTAATTTATTAATTTATTAATTTATTAATTTATTAATTTA 1005
QY 2200 CCTACTTAAAGATGAGAACCAATAGATATCTATCTTAAGTATGATATCCCTACGAT 2259
DB 1006 TATTAATTTATTAATTTATTAATTTATTAATTTATTAATTTATTAATTTATTAATTTA 1065
QY 2260 GTATCAATTAATTAAGAAAAATGTTATGATATTAATTAATTAATTAATTAATTTCAACT 2319
DB 1066 ATATTTATTAATTTATTAATTTATTAATTTATTAATTTATTAATTTATTAATTTATTT 1125
QY 2320 CACACAAACACCAAGTCTATTAATTAACAAATTC--AGAAATTAATTAATTTACCA 2377
DB 1126 TTAATTTATTAATTTATTAATTTATTAATTTATTAATTTATTAATTTATTAATTTA 1185
QY 2378 ACACATAGTATCAATTTCTATTAATTAATTAATTTATTAATTTATTAATTTATTAATTTA 2437
DB 1186 TTAATTTATTAATTTATTAATTTATTAATTTATTAATTTATTAATTTATTAATTTA 1245
QY 2438 TCATCATTAATTAATTAATTAATTAATTAATTTATTAATTTATTAATTTATTAATTTA 2497
DB 1246 AAAAAAAAAAANN 1305
QY 2498 TTCAATTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2557
DB 1306 TATTAATTTATTAATTTATTAATTTATTAATTTATTAATTTATTAATTTATTAATTTA 1365
QY 2558 TTACTCCAAACAGATTAATTAATTTATTAATTTATTAATTTATTAATTTATTAATTTA 2596
DB 1366 TTAATTTATTAATTTATTAATTTATTTTATTTATTAATTTATTAATTTATTAATTTA 1404

RESULT 30

EZ696936/c 1074 bp DNA linear GSS 02-JUL-2003
LOCUS
DEFINITION
SP_Ba0091G12.f SP_Ba Sorghum propinquum genomic clone
SF_Ba0091G12 5', genomic survey sequence.
E2696936
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Sorghum propinquum
Sorghum propinquum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 1074)
Wing, R., Yu, Y., Kim, H.R., Collura, K., Pries, G., Currie, J.,
Soderlund, C., and Hatfield, J.
Sequencing of Sorghum BAC ends.
http://genome.arizona.edu/stc/sorghum
Unpublished (2003)
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu

REFERENCE

Wing, R., Yu, Y., Kim, H.R., Collura, K., Pries, G., Currie, J.,
Soderlund, C., and Hatfield, J.
Sequencing of Sorghum BAC ends.
http://genome.arizona.edu/stc/sorghum
Unpublished (2003)
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu

TITLE

JOURNAL
COMMENT

FEATURES

source
Location/Qualifiers
1..1074
/organism="Sorghum propinquum"
/mol_type="genomic DNA"

/db xref="taxon:132711"
 /clone="SP_Ba091G12"
 /note="Vector: pBelBAC11, Site_1: HindIII, Site_2:
 HindIII; Paterson lab BAC library (HindIII)"

ORIGIN

Query Match 2.5%; Score 119.4; DB 28; Length 1074;
 Best Local Similarity 44.6%; Pred. No. 1.3e-08;
 Matches 471; Conservative 0; Mismatches 576; Indels 8; Gaps 1;

1412 TTGGACTAGAAATAGAGAAAGTAGAAATATTAATACATTATAGATCAACCTTAAG 1471
 1058 TATATAATAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATAG 999
 1472 CTGTGTGAAGACAGACTGTAAACCACTGATATCATTAGAAATTAATTAATGTA 1531
 998 AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 939
 1532 GAATACCAATTTGCTATCTATTACATTTTATTAATGAGATATATGATGATTAAT 1591
 938 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 879
 1352 GCCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1651
 878 AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 819
 1652 TAAACGGTCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1711
 818 AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 759
 1712 CATGTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1763
 758 AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 699
 1764 TAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1823
 698 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 639
 1824 AATGTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1883
 638 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 579
 1884 TTGGACACTTTTCACTACATCTATTAATTAATTAATTAATTAATTAATTAATTA 1943
 578 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 519
 1944 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2003
 518 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 459
 2004 ATAGTGTCTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2063
 458 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 399
 2064 TAAATGTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2123
 398 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 339
 2124 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2183
 338 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 279
 2184 AATTTCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2243
 278 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 219
 2244 GATATCCTTACAGATGATCAATTAATTAATTAATTAATTAATTAATTAATTAAT 2303
 218 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 159
 2304 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2363

Db 158 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 99
 Qy 2364 TATATACCTTCCACACATAGTATTAATTAATTAATTAATTAATTAATTAATTA 2423
 Db 98 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 39
 Qy 2424 GATGACTTCCACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2458
 Db 38 NNN 4

RESULT 31
 CDS00EVL 1101 bp DNA linear GSS 04-JUN-1999
 LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC:
 DEFINITION BACR29B23 of RPCT-98 library from Drosophila melanogaster (fruit
 fly), genomic survey sequence.

ACCESSION AL069706.1 GI:4949849
 VERSION AL069706
 KEYWORDS GSS.
 SOURCE Drosophila melanogaster (fruit fly)
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Preygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 1101)
 Genoscope.
 Direct Submission
 Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
 BP 191 91006 Evry cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)

COMMENT
 Determination of this BAC-end sequence was carried out as part of a
 collaboration with the Berkeley Drosophila Genome Project (BDGP).
 The BDGP is constructing a physical map of the Drosophila
 melanogaster genome using these BACs. For further information
 please see <http://www.fruitfly.org> The BDGP Drosophila
 melanogaster BAC library was prepared by Kazuo Osoegawa and
 Aaron Mammeter in Pieter de Jong's laboratory in the Department of
 Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
 NY. The library is named RPCT-98 and was constructed by partial
 EcoRI digestion of Drosophila DNA provided by the BDGP from the
 isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
 p1 and EST libraries. A more detailed description of the library
 and how to order individual BAC clones the entire library, or
 filters for hybridization from the BACpac Resource Center can be
 found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
 source
 1..1101
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"
 /db_xref="taxon:7227"
 /clone="BACR29B23"
 /clone_11b="RPCT-98"
 /note="end : 17"

ORIGIN

Query Match 2.5%; Score 119.4; DB 29; Length 1101;
 Best Local Similarity 36.0%; Pred. No. 1.3e-08;
 Matches 237; Conservative 131; Mismatches 285; Indels 2; Gaps 2;

1849 CAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1908
 443 NNN 502
 1909 GATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1968
 503 NNN 562
 1969 AATTTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2028
 563 NNTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 622
 2029 TATCATCTTGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2088

[illegible]

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	PUBMED	COMMENT
CG749971.c	CG749971	P044-2-E09.za	Ppa	ECORI	BAC Library	<i>Pristionchus pacificus</i> genomic.	1377 bp	DNA	linear	GSS 24-OCT-2001			
						genomic survey sequence.							
		CG749971											
		CG749971.1	GI:37970948			GSS.							
						<i>Pristionchus pacificus</i>							
						<i>Pristionchus pacificus</i>							
						<i>Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;</i>							
						<i>Neodiplogasteridae; Pristionchus</i> .							
						1 (bases 1 to 1377)							
						Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perbolte,L., Jansen,K.,							
						Bunjer,J., van der Meulen,M. and Sommer,R.J.							
						An integrated physical and genetic map of the nematode <i>Pristionchus</i>							
						<i>pacificus</i>							
						Mol. Genet. Genomics 269 (5), 715-722 (2003)							
						22835951							
						12884007							
						Contact: Sommer R.J							
						Evolutionary Biology							
						Max-Planck-Institute for Developmental Biology							
						Spemannstr. 37-39, Tuebingen D-72076, Germany							
						Tel.: 00497071601371							
						Fax: 00497071601498							
						Email: ralf.sommer@tuebingen.mpg.de							
						Class: BAC ends.							

FEATURES	Location/Qualifiers
source	1. .1377

```
/organism="Pristionchus pacificus"  
/mol_type="genomic DNA"
```

```
/strain="California"  
/db_xref="taxon:54126"
```

```
/clone_lib="Ppa EcoRI BAC Library"  
/note="The library was generated by a partial digest of
```

the genomic DNA with EcoRI and cloning into the BAC vector. "

Query Match 2.5%; Score 119.4; DB 29; Length 1377

Best Local Similarity 40.6%; Pred. NO. 1.2e-08;
Matches 498; Conservative 0; Mismatches 726; Indels 2; Gaps 1;

QY	1274	ATACTAAAGCATGATGAAATTTAAACACTATATATATAGCAAAATATCATTCATATACAT	1333
Db	1358	AA	1299
QY	1334	CTTCATACATTAAATGGTCATGTCGTTGGAGACTCTATATTTAACACTGACACAGAA	1399
Db	1298	NNNNNANNNAAA	1239
QY	1394	CATTCACGCATATATGGCTTTTGGACTGAAATAGGAAAGGTAGAAATATATATACATTA	1453
Db	1238	AA	1179
QY	1454	ATGATGACAACTCTAAAGCTGTGTTGAAAGAACAGTACTGTAAACCACTGATATCATTA	1513
Db	1178	AA	1119
QY	1514	GCAGAAAACTTATGTTAGAACTACATTTGCATATCTATTTCAATTTATTTACATGAG	1579
Db	1118	AA	1059
QY	1574	AATATATCCTAGATTAAGCCATATCAACATATTAATCTTAGTCTATCATCTTGGATAA	1633
Db	1058	AA	999
QY	1634	TTTCACACATGATTAATTAACGGTCAGAAATATAGAAATCAACGTTCAACCATATAT	1699
Db	998	AA	939
QY	1694	AAAAAGCATTAATCCCATCGTTAACATCACTAACTAAGTATGTACATATTAACATACA	1753
Db	938	AA	879
QY	1754	AACCTTAAGTATACCTTAAATTTCTCTAGCTATATATATAACTAACTCTATCTCTATC	1813
Db	878	AA	819
QY	1814	TCAATTAACCATGTGTATCAGATTATTAATTAAGTCAATTAATACTTTACTTATTTATA	1879
Db	818	AA	759
QY	1874	AAATATATGTGTGCAACTTTTTCATCACTCATGTATATATGACATATATTTACAT	1933
Db	758	AA	699
QY	1934	AAATTACTTAATAAAAGAACCAATACATATATTAATTAATTAATAAACAACCTATTACTAC	1993
Db	698	AA	639
QY	1994	GTAATATTACATGATGTCCTGACATTAATAATTCTATCATCTTGACATATTTCAACA	2053
Db	638	AA	579
QY	2054	TAAATATGATAAATCTGTTTTTATATGCCATTAGTTCTTATATAAATTTGCTAATTACC	2113
Db	578	AA	519
QY	2114	AAAAAGTTAAATACACATTTCTGCAACAATATGAGAGAGATCAATACACGTCAAATA	2173
Db	518	AA	459
QY	2174	AATTTAAACAATTTCTAATATACACCACTACTTAACATAGGAACCATGTGATATCAT	2233
Db	458	AA	399
QY	2234	ACCTAAGTATATATCCTTACAGCATGTATCAATAATTAAGAAATAATCGTTATGTATAT	2293
Db	398	AA	339
QY	2294	TTATATCAATTTCAAAATTTCCAACTTCAACACAAATCTCAAAGTTCTATATATACAAAT	2353

Db 338 AAAAAAAAAAANNANNTA--ACATACATACATNACTCTCTCTTAACATA 281

QY 2354 TGAAGATTATATATCTTACCAACACATGATTAATCTTAATACATATCC 2413

Db 280 CCAATTTACCCATATNCCACCTTANTTTCTCTCTCAATCAAAAANTTAATCA 221

QY 2414 TTTTAAAGATGACTTCAATCATATCAATTAATTAATGATGCAATGTTATCA 2473

Db 220 TCAACCATCTCCTNNTCTNCAATCTTATNATACANCTNCAACAACTNCAANTN 161

QY 2474 GAACATTCACATCAATCAAGATTTT 2499

Db 160 TNAACCTCAAAATATTTCAATATAT 135

RESULT 33
BX436282/c 1124 bp mRNA linear EST 15-MAY-2003

LOCUS BX436282 Homo sapiens THYMUS Homo sapiens cDNA CSOCAP001YC01

DEFINITION 5-PRIME, mRNA sequence.

ACCESSION BX436282.1 GI:30787521

VERSION EST.

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 1124)

AUTHORS Li W.B., Gruber C., Jessee J. and Polayes D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished (2001)

COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: segre@genoscope.cns.fr; Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of Invitrogen. Contact: Feng Liang Email: fliang@lifetech.com URL: http://fulllength.invitrogen.com/Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID: CSOCAP001AB010P1.

FEATURES
source location/Qualifiers
1..1124
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSOCAP001YC01"
/cissue_type="THYMUS"
/clone_lib="Homo sapiens THYMUS"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was enriched with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN

Query Match 2.5%; Score 117.8; DB 13; Length 1124;
Best Local Similarity 27.1%; Pred. No. 2.2e-08;
Matches 215; Conservative 255; Mismatches 308; Indels 16; Gaps 2;

QY 1564 ATTACATGAAATATATCAATGATTAAGCATATCAACATTAATCTTATGCTATCAT 1623

Db 1121 AAAAAAAAAAHH 1062

QY 1624 CTTTGATATTTTCAACATGATTAATTAAGCTCAAAATATAGATGATCAACGTTCA 1663

Db 1061 YNYYHHHAAAYNYYHH 1002

QY 1664 CCACATATATTAAGAGCATGCAATATCCATGTTAATCAATCAACTAAGATATGTTACATA 1743

Db 1001 YYYYYYAAAAAAAAAAAAAAAAHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH 942

QY 1744 TTACATACAACTTAATGATTAATTAATTTCTCTGATATATATTAACAAATTCCT 1803

Db 941 YYYYYYHH 882

QY 1804 ATCTCTATCTCAATTAACCAATTTGTTATTCAGTTATTAATAAGTC-----A 1850

Db 881 YYYYYHHHHAA 822

QY 1851 AATTAACCTTTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1910

Db 821 AA 762

QY 1911 ATATATGAACATTAATTTTATGATTAATACTATATTAATAAGAACCAATCATATTA 1970

Db 761 MHHYAAAAAAAAAAAAAAAAHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH 702

QY 1971 TTACTAAACACCATTTTACTAGTAATTAATTAATTAATTAATTAATTAATTAATTA 2030

Db 701 YYYYYYHH 642

QY 2031 TCATCTTGACATATTTTCAACATTAATAATGATTAAGCTGTTTATGCCCCATTA 2090

Db 641 YYYYYYHH 582

QY 2091 CTTAATTAATTTGCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2150

Db 581 YYYYYYHH 522

QY 2151 GAAGATTAACATACAGCTCAATTAATTAACAAATTTCTAATTAATTAACCACTCTA 2210

Db 521 AAAAAAAAAAAAAA---HAAAAAAAAAAAYAAHHHHHHHHHHHHHHHHHHHHHH 465

QY 2211 CATAGAACCATATGATATCATATCTTAAGTAAAGATATCCCTACAGCATGATCA 2270

Db 464 AAAAAAAAAAAAYHH 405

QY 2271 TAAAGAAAAATCGTATGATATTTATTAATTAATTAATTAATTAATTAATTAATTA 2330

Db 404 YHH 345

QY 2331 CTCAAGTCTTAT 2344

Db 344 YYYAAHH 331

RESULT 34
CC231597/c 1426 bp DNA linear GSS 12-MAY-2003

LOCUS CH261-36A19.RM1.1 CH261 Gallus gallus genomic clone CH261-36A19,

DEFINITION genomic survey sequence.

ACCESSION CC231597

VERSION CC231597.1 GI:30558260

KEYWORDS GSS.

SOURCE Gallus gallus (chicken)

ORGANISM Gallus gallus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.

AUTHORS Krenitzki, C., Higginbotham, J., Wylie, K., Carter, J., McPherson, J., Warren, W., Graves, T., Mardis, E. and Wilson, R.

TITLE Gallus gallus BAC End Reads

JOURNAL Unpublished (2003)

COMMENT Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@wustl.wustl.edu
Insert Length: 18200 Std Error: 0.00
Seq primer: RM1 TACGACTCATATAGGAGAGA
Class: BAC ends
High quality sequence start: 276
High quality sequence stop: 339.

FEATURES
source location/Qualifiers
1..1426
/organism="Gallus gallus"
/mol_type="genomic DNA"

QY 1514 GCAAAAACCTTAATAGAACCTACCAATTGTCATATCTATTACATTTATTACATAG 1573
 DB 313 AA 372
 QY 1574 AATATATCATCTAGATTAAGCCATATCAACATATTAATCTTATCTATCATCTTTGATTA 1633
 DB 373 AA 432
 QY 1634 TTTCACACATGATTAATTAACGGTCAAAATATAGATTAACAGTTTCAACATATAT 1693
 DB 433 AGAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 492
 QY 1694 AAAAAAGCATCAATATCCCATGTTAATCACTCAAGTATGATTAATTAATTAACA 1753
 DB 493 AA 552
 QY 1754 AACTTAAGATTAATTAATTTCTAGTATTTATTAACAAATTTCTATCTCTATC 1813
 DB 553 AACAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 612
 QY 1814 TCATTAACCAATTTGTTATCAGATTATTAATTAAGTCAATTAATTTATCTTATTTATA 1873
 DB 613 AA 672
 QY 1874 AATTAATATGTTGCACTTTTTCATCATCATGATATATATGATTAATTAATTTTACAT 1933
 DB 673 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 732
 QY 1934 AAAAAATCACTTAATTAAGAACATATATTAATTAATTAATTAATTAATTAATTAAT 1993
 DB 733 AA 792
 QY 1994 GTAATATTAATGATGCTGTCATATTAATTTATCTATCTTTGATTAATTTTCAACA 2053
 DB 793 AA 852
 QY 2054 TAAATATATGATTAATGCTTTTATGCCATTAATGCTTATTAATTAATTTGATTAAT 2113
 DB 853 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 912
 QY 2114 AAAAAAGTTAAATTAATCACTCTGCAATATTAATTAAGTATTAATTAATTAATTAAT 2173
 DB 913 AA 972
 QY 2174 AATTAACCAATTTCTATATATCACTCTTAACATTAAGAACATTAATTAATTAAT 2233
 DB 973 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 1032
 QY 2234 ACCTAAGTATGATTCCTGACATGATTAATTAATTAAGAACATTAATTAATTAAT 2293
 DB 1033 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 1092
 QY 2294 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2353
 DB 1093 AA 1152
 QY 2354 TCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2413
 DB 1153 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 1212
 QY 2414 TTTTAAAGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2456
 DB 1213 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 1255

RESULT 36

CNS033GQ

Tetradon nigroviridis genome survey sequence PUC-Ort end of clone

20824 of 118741 G from Tetradon nigroviridis, genomic survey

ACCESSION AL226115.1 GI:7885026

KEYWORDS

GSS: genome survey sequence.
Tetradon nigroviridis

SOURCE

Tetradon nigroviridis

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes; Tetraodontidae; Tetraodontidae; Tetraodon.

REFERENCE

1. Roest Crolius, H., Jallion, O., Daeilva, C., Bouneau, L., Fisher, C., Bernot, A., Fzames, C., Wincker, P., Brothier, P., Quetier, F., Saurin, W., and Weissenbach, J.

AUTHORS

Roest Crolius, H., Jallion, O., Daeilva, C., Bouneau, L., Fisher, C., Bernot, A., Fzames, C., Wincker, P., Brothier, P., Quetier, F., Saurin, W., and Weissenbach, J.

TITLE

Estimate of human gene number provided by genome-wide analysis using Tetradon nigroviridis DNA sequence

JOURNAL

Nat. Genet. 25 (2), 235-238 (2000)

MEDLINE

20296633

PUBMED

10835645

REFERENCE

2. Roest Crolius, H., Jallion, O., Daeilva, C., Ozouf-Coataz, C., Fzames, C., Fischer, C., Bouneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A., and Weissenbach, J.

AUTHORS

Roest Crolius, H., Jallion, O., Daeilva, C., Ozouf-Coataz, C., Fzames, C., Fischer, C., Bouneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A., and Weissenbach, J.

TITLE

Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetradon nigroviridis

JOURNAL

Genome Res. 10 (7), 939-949 (2000)

MEDLINE

20358837

PUBMED

10899143

REFERENCE

3. (bases 1 to 1135)
Submitted (12-APR-2000) Genoscope - Centre National de Sequencage; BP 191 91006 EVRY cedex - FRANCE (E-mail: seqref@genoscope.cns.fr)

COMMENT

This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetradon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/Tetradon>.

FEATURES

Location/Qualifiers
1..1135
/organism="Tetradon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone="20824"
/clone_1ib="G"
/note="Genoscope sequence ID : CNS033GQ1251-end : PUC-Ort"

ORIGIN

Query Match 2.5%; Score 117; DB 29; Length 1135;
Best Local Similarity 45.5%; Pred. No. 2.9e-08;
Matches 429; Conservative 15; Mismatches 486; Indels 10; Gaps 2;

QY 1507 ATCAATTAAGCAAAAACCTTAATGATTAAGCAATTTGCAATTAATTTTACATTTATT 1566
 DB 199 ATATATAA 258
 QY 1567 ACATGAATATATCACTGATTAAGCAATTAATTAATTAATTAATTAATTAATTAAT 1626
 DB 259 AA 318
 QY 1627 TTGATTAATTTCAACACATGATTAATAACGGTCAAAATTAATTAATTAATTAATTAAT 1686
 DB 319 AA 378
 QY 1687 CATATTAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1746
 DB 379 AA 438
 QY 1747 CATTAACCACTTAATTAATTAATTAATTTCTAGTAAATTAATTAATTAATTAATTTCTATC 1806
 DB 439 AA 498
 QY 1807 TCTATCTCAATTAACCAATTTGTTATCAGATTAATTAATTAATTAATTAATTAATTTTACT 1866
 DB 499 AATTAATAA 558

QY 1867 ATTTAAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 1926
DB 559 ATTTAAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 616
QY 1927 TTACATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 1986
DB 617 TTACATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 676
QY 1987 TTACATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 2046
DB 677 TTACATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 736
QY 2047 TTACATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 2106
DB 737 TTACATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 796
QY 2107 TTACATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 2166
DB 797 TTACATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 856
QY 2167 TTACATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 2226
DB 857 TTACATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 916
QY 2227 TTACATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 2286
DB 917 TTACATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 968
QY 2287 TTACATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 2346
DB 969 TTACATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 1028
QY 2347 TTACATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 2406
DB 1029 TTACATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 1088
QY 2407 TTACATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 2448
DB 1089 TTACATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 1130

RESULT 37
CG757211/c 1269 bp DNA linear GSS 24-OCT-2003
LOCUS P052-2-005.2a Ppa EcORI BAC Library Pristionchus pacificus genomic,
DEFINITION genomic survey sequence.
ACCESSION CG757211
VERSION CG757211.1 GI:37985546
KEYWORDS GSS.
SOURCE Pristionchus pacificus
ORGANISM Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
Neodiplogasteridae; Pristionchus.
1 (bases 1 to 1269)
AUTHORS Srinivasan,J., Stenz,W., Jesse,T., Wieggers-Perebole,L., Jansen,K.,
Bunjer,J., van der Meulen,W. and Sommer,R.J.
TITLE An integrated physical and genetic map of the nematode Pristionchus
pacificus
JOURNAL Mol. Genet. Genomics 269 (5), 715-722 (2003)
MEDLINE 22835951
COMMENT Contact: Sommer RJ
12884007
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
Class: BAC ends.
FEATURES
source
1..1269
/organism="Pristionchus pacificus"

ORIGIN
Query Match 2.5%; Score 117; DB 29; Length 1269;
Best Local Similarity 41.6%; Pred. No. 2.8e-08;
Matches 429; Conservative 0; Mismatches 602; Indels 0; Gaps 0;
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Ppa EcORI BAC Library"
/note="The library was generated by a partial digest of
the genomic DNA with EcORI and cloning into the BAC
vector."

QY 1373 ATATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 1432
DB 1267 ATATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 1208
QY 1433 ATATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 1492
DB 1207 ATATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 1148
QY 1493 ATATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 1552
DB 1147 ATATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 1088
QY 1553 ATATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 1612
DB 1087 ATATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 1028
QY 1613 ATATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 1672
DB 1027 ATATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 968
QY 1673 ATATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 1732
DB 967 ATATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 908
QY 1733 ATATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 1792
DB 907 ATATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 848
QY 1793 ATATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 1852
DB 847 ATATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 788
QY 1853 ATATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 1912
DB 787 ATATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 728
QY 1913 ATATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 1972
DB 727 ATATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 668
QY 1973 ATATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 2032
DB 667 ATATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 608
QY 2033 ATATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 2092
DB 607 ATATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 548
QY 2093 ATATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 2152
DB 547 ATATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 488
QY 2153 ATATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 2212
DB 487 ATATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 428
QY 2213 ATATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 2272
DB 427 ATATTAATTAATGTTTGCACCTTTTCACTACATGATATATATAGACATAT 368

COMMENT

Contact: Genoscope
Genoscope - Centre National de Séquençage
BP 191 91006 Evry cedex - France
Email: segref@genoscope.cns.fr Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 7009.f For
more information about this cluster, see
[http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0CAP004DD02QPI&cluster=7009.f](http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0CAP004DD02QPI&cluster=7009.f). Contact
Peng Liang Email : liang@life.techn.com URL :
<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0CAP004DD02QPI.

ORIGIN

Query Match	2.4%;	Score 115.8;	DB 13;	Length 1061;
Best Local Similarity	42.4%;	Pred. No. 4.5e-08;		
Matches 297;	Conservative	56;	Mismatches 343;	Indels 4;
			Gaps	12

```

/clone_lib="Homo sapiens THYMUS"
/notes="vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

```

D_b 933 YYYTYTYYTTATTAATAAAAAAAYAAAAAYAAACAAAACAAAAAACCACTAT 998
D_b 933 YYYTYTYYTTATTAATAAAAAAAYAAAAAYAAACAAAACAAAAAACCACTAT 998
Q_y 2157 ACATATCAGCTCAATTAAATTAAACAATTTCTAATATPA 2196
Q_y 2157 ACATATCAGCTCAATTAAATTAAACAATTTCTAATATPA 2196
D_b 999 ACATATCTAATACGTAACAAAAACAAAAAYATGTAAACVMAA 1038
D_b 999 ACATATCTAATACGTAACAAAAACAAAAAYATGTAAACVMAA 1038

RESULT 40
CG744200/

LOCUS	CG744200	1305 bp	DNA	linear	GSS 24-OCT-2003
DEFINITION	P036-3-H01.za Ppa ECORI BAC Library Pristionchus pacificus genomic, genomic survey sequence.				

FEATURES

```

source
1.1305
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="pca EcoRI BAC Library"
/note="The library was generated by a partial digest of
the genomic DNA with EcoRI and cloning into the BAC
vector."

```

ORIGIN

Query Match	2.4%	Score 115.6;	DB 29;	Length 1305;
Best Local Similarity	42.6%;	Pred. No. 4.5e-08;		
Matches 379;	Conservative 0;	Mismatches 511;	Indels 0;	Gaps 0;

QY	1420	ACAAATAGAAAAGGTGAAATTAATACTTAATAGATCAACCTTAAGCTGTGGA	147
Db	1291	AA	1238
QY	1480	AGGAACGAGTCTGTGTAACCACTGATATCATGACAAAAAACTTAATGTGAAGTACG	1538
Db	1231	AA	1178
QY	1540	AATTGGCATCTATCTTTTCACATTTATTACATGAGATATATCACTAGATTAAGCCATATC	1599
Db	1171	AA	1112
QY	1600	AAACATATTATCTTAGTCTATCATCTCTTGATTAATTCAACAATGATTAATAATAAACGT	1655
Db	1111	AA	1052
QY	1660	CAAAAATATAGATACACAGCTTCGCCACATATTAATAAAGCAGTCAAAATACCCATCGTTA	1719
Db	1051	AA	992
QY	1720	ACATCCAACTAGATGTAACTATTTACATACAAACCTTAAGTATTAATTAATAATTTCTCT	1779

